Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Effect of First-Order Conditional Probability in a Two-Choice Learning Situation: N. H. ANDERSON 73

Relation of Serial Position Errors to Doublet and Split-Doublet Location in Verbal Maze Pattern: R. L. ERNST, D. R. HOFFELD, S. SEIDENSTEIN, AND W. J. BROGDEN 94

Retroactive Inhibition of Connected Discourse as a Function of Practice Level: N. J. SLAMECKA 104

Motor Skill Transfer as a Function of Intertask Interval and Pretransfer Task Difficulty: G. NAMIKAS AND E. J. ARCHER 109

Anchor, Contrast, and Paradoxical Distance Effects: H. HELSON AND M. C. NASH 113

Eliminating a Conditioned GSR by the Reduction of Experimental Anxiety: R. E. SILVERMAN 122

Partial and Correlated Reward in Escape Learning: G. H. BOWER 126

The Effect of Stimulus Similarity on Amount of Cue-Position Patterning in Discrimination Problems: B. N. WHITE AND C. C. SPIKER 131

American Psychological Association

Vol. 59, No. 2 February 1960

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Address all articles submitted for publication to the Editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. **Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association.*** When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1960 the publication cost is $30 per printed page.

To save printing costs, supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
BEATRICE B. HAMLIN, Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1960 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 59, NO. 2 FEBRUARY 1960

EFFECT OF FIRST-ORDER CONDITIONAL PROBABILITY IN A TWO-CHOICE LEARNING SITUATION [1]

NORMAN H. ANDERSON [2]

University of Wisconsin

The experiments reported here investigate behavior in a two-choice probabilistic task in which the independent variable is the conditional probability in the sequence of stimulus events. Each of the two stimulus events (lights) was required to occur equally often, but a partial patterning was introduced by varying the first-order conditional probability in the sequence of lights. This variable, denoted by π_{11}, is defined as the probability of occurrence of Light 1 on any trial, given that Light 1 occurred the previous trial. When $\pi_{11} < .5$, the lights tend to alternate on successive trials; when $\pi_{11} > .5$, each light tends to repeat itself. Thus, $\pi_{11} = .7$ represents a 70% repetition pattern, and $\pi_{11} = .3$ represents an equally strong alternation pattern. Since each light occurs equally often, the appropriate dependent variable is the frequency of *repetition responses*, a repetition response being defined as the response of predicting on the current trial that same *light* which occurred on the previous trial. An alternation response is analogously defined as predicting that light which did not occur on the previous trial.

This conditional probability variable was first used by Hake and Hyman (1953). It is of considerable interest because of its close relationship to the *frequency* variable for which the Estes and Straughan (1954) model predicts the "matching" behavior first reported by Grant, Hake, and Hornseth (1951). Engler (1958) tested a Bush-Mosteller (1955) model which predicts both the absolute frequencies of the two choices, and the level of repetition responses. The model succeeded with the first dependent variable, but failed to fit the repetition response results.

The first of the experiments reported here was designed to test a model (Anderson, 1956) for the con-

[1] Based on work submitted in partial fulfillment of the requirements for the PhD degreee at the University of Wisconsin, 1956. The work was done while the author was an NSF predoctoral fellow. I wish to express my thanks to D. A. Grant, W. J. Brogden, and H. F. Harlow for their advice and helpful criticism. Thanks are also due S. H. Sternberg and R. C. Atkinson for a critical reading of the manuscript.

[2] Now at the University of California, Los Angeles.

ditional probability situation. The model is based on the ideas of Estes (1950) and of Restle (1955). It is assumed that there are two distinct subsets of stimulus elements, one of which is specific to the repetition response, the other specific to the alternation response. On any trial the sampled stimulus elements specific to a given response become conditioned to that response if it was indicated as correct, regardless of whether S actually made the response. However, if the given response was not the correct one, then the sampled stimulus elements specific to that response go into a neutral state. The probability of a repetition response on any trial is the ratio of the number of sample elements conditioned to the repetition response to the total number of sample elements in the conditioned state. Neutral elements have no effect on response probability, but they may be conditioned on later trials. (The model also postulated a third set of stimulus elements corresponding to responses other than the two above. It was assumed that such stimulus elements were progressively neutralized over trials. This third set of elements would thus affect the learning rate, but not the asymptotic level. Hence, it need not be further considered here.)

This model leads to the following approximate expression for the asymptotic probability of repetition responses:

$$\frac{1}{1 + c(1 - \pi_{11})/\pi_{11}} \qquad [1]$$

where π_{11} is the conditional probability of Light 1, given that Light 1 occurred on the previous trial, and c is a constant which measures the relative weights of the two sets of stimulus elements.

Since the completion of the present work, Burke and Estes (1957) have revised the original Estes and Straughan (1954) model, taking into account the nonindependence of the sample of active stimulus elements on successive trials. The revision, it may be noted, yields some small changes in the predicted asymptotes for the event frequency situation. More pertinently, it also yields the following expression for asymptotic repetition response level:

$$\pi_{11} + \theta\pi_{11}(1 - \pi_{11}) \qquad [2]$$

where θ is the learning rate. This model will also be tested against the present data.

Upon completion of Exp. I, it was seen that the model, although it gave a reasonable account of the asymptotic acquisition performance, was unable to handle the behavior in the transfer condition in which the various acquisition groups failed to converge to a common level despite the fact that they were receiving the same treatment. Since this differential transfer effect appeared to create difficulties for the mathematical approach in general, it seemed desirable to abandon the testing of specific models and begin a more extensive empirical investigation. Experiment II was accordingly designed to give a parametric treatment of the conditional probability variable. The data were recorded so as to permit the easiest tabulation of the sequential dependencies in the stimulus-response sequences. It was expected, both on theoretical grounds (Anderson, 1959) and from the results of earlier work (Estes & Straughan, 1954; Anderson & Grant: 1957, 1958; Jarvik, 1951), that these dependencies would give additional insight into the processes underlying the behavior in this situation.

METHOD

The procedure common to the two experiments will be given first; that particular to the individual experiments will be found under the corresponding headings. Each group is labeled by 100 times its acquisition value of π_{11}. Thus, Group 30 received $\pi_{11} = .3$ in acquisition, etc.

Apparatus

The apparatus was the same as that used in previous work (Anderson & Grant, 1957). Two 6-w. opal lights, mounted 10 in. apart horizontally on a vertical, flat black, 15 × 20-in. panel, formed a stimulus display. The two lights were activated by a punched tape fed through a Western Union tape transmitter. An 800-cps tone served as the signal to respond. Presentation of tone and lights was automatic.

A row of four partially separated booths was placed 9 ft. from the panel of lights. Each booth contained a small box with two non-self-releasing keys, corresponding to the two lights. The *S*s were required to push the keys down, and then push them up, in making their responses. Stimuli and responses were recorded automatically with an Esterline-Angus operation recorder.

Except for the panel of lights and the booths, all equipment was placed on *E*'s desk which lay behind the booths, and from which *E* monitored the experiment.

Procedure

The *S*s were volunteers from elementary psychology classes who received class points for their work. They were randomly assigned, in groups of two, three, or four, to the various experimental conditions.

At the beginning of the experiment, *S*s were seated in the booths and read the instructions which were similar to Instructions C of Anderson and Grant (1957). In addition to operating detail concerning the keys, tone, and lights, the instructions stated that the task was to predict which of the two panel lights would come on next, that this was an experiment in learning, and that *S*s should do the best they could even though the task seemed difficult. Three practice trials, in which both lights flashed each time, were then given, and questions answered. Finally, *S*s were told that the rest of the experiment would have to be run off without conversation or interruption. The sequence of trials was continuous, with no external indication of change from acquisition to transfer condition.

Independent Variables

The two experimental variables were trial rate, and first-order conditional probability. The temporal division of each trial was as follows: signal tone (during which responses were to be made), 2.0 sec.; waiting time, 1.2 sec.; stimulus light on, .8 sec.; "intertrial interval," 1.0, 3.5, or 6.0 sec. These intertrial intervals thus yielded trial rates of 5.0, 7.5, and 10.0 sec. per trial.

First-order conditional probability has been defined in the introduction. The sequences of lights were random subject to the following rules: (*a*) exactly one light flash each trial; (*b*) the assigned value of π_{11} hold exactly within each block of 50 trials; (*c*) the absolute frequency of each light lie between 20 and 30 in each block of 50 trials. Four independent sequences were constructed for each acquisition value of π_{11}. In the transfer condition, four independent sequences were also used, but these were the same for all groups within each experiment.

Experiment I.—This experiment was based on a 3 × 2 design using the three trial rates, 5, 7.5, and 10 sec. per trial, and the two acquisition values of conditional probability, $\pi_{11} = .3$, and $\pi_{11} = .7$. Light sequences were kept orthogonal to these two variables. Each group received 200 acquisition trials. In the transfer condition, $\pi_{11} = .5$ for all groups. The 5, 7.5, and 10 sec. groups were given 300, 200 and 100 transfer trials, respectively.

Experiment II.—This experiment was a parametric study of the effect of conditional probability on acquisition, on transfer to an extinction condition, and on recovery. A group of *S*s was run at each of the 11 acquisition values of π_{11} ranging from 0 to 1 in steps of .1. The transfer condition was the same for all groups, with $\pi_{11} = .5$ as in Exp. I. Acquisition and transfer were given in Session 1; recovery was tested in Session 2.

Trials were presented at the 5-sec. rate used in Exp. I. New stimulus sequences were constructed for conditions duplicating those of Exp. I. Instructions were reworded slightly by incorporating an additional statement to the effect that, while *S* should be able to do considerably better than just guessing, a perfect score was impossible, so that *S* should not worry if he was not correct every time.

Groups 10–90 were given 300 acquisition trials, followed by 200 transfer trials. Groups

0 and 100 received 100 acquisition trials, and 400 transfer trials. For these two groups, right and left were counterbalanced in Trials 1–100 within each of the four transfer sequences. Twenty *Ss* were run in each group, except that Group 50 contained 40 *Ss*.

Two additional subgroups, each with $N = 12$, were run with $\pi_{11} = .3$, and .7. The sequences for these groups were identical with those of the corresponding conditions of Exp. I for the first 200 trials, and with those of Exp. II for the remaining trials. These two groups permitted an evaluation of effect of the change in instructions from Exp. I to Exp. II, without confounding any possible effect of sequences. Since no effect of instructions was found, these *Ss* have been combined with the others so that Groups 30 and 70 each have $N = 32$.

Recovery was tested by giving 60 additional trials with $\pi_{11} = .5$. Fifteen such sequences were constructed by taking successive blocks of 60 trials from three of the transfer sequences of Exp I. The *Ss* were assigned to these sequences unsystematically, but each sequence was used about equally often. Session 2 was run after an interval of two to six days. An *S* was signed up for Session 2 before serving in Session 1, according to *E*'s schedule and each *S*'s own convenience. In contrast to Session 1, *Ss* were occasionally run alone in Session 2.

One *S* adopted the deliberate policy of leaving his key down for several trials until he wished to change his choice of responses. He was replaced although his record seemed normal in other respects. Three *Ss*, all from the additional subgroup run under the .7 condition, did not return for Session 2. The Session 2 record of one *S* was lost because of a recording error.

Results

A repetition response is defined as a press of that key corresponding to the light which flashed the previous trial. Since two stimuli and two responses are involved, a nonrepetition response may be called an alternation response. For conditions with $\pi_{11} > .5$, a repetition response is also an optimal response since it is more likely to be correct. Similarly, when $\pi_{11} < .5$, an alternation response is an optimal response. The main dependent variable is repetition response frequency.[3]

Unless otherwise noted, the analysis of variance was used in the statistical analyses. The Sequence variable, corresponding to the different random sequences used, was treated throughout as a fixed variate (Wilk & Kempthorne, 1955, p. 1163).

Mean Performance Curves

Experiment I.—The results of Exp. I are shown in Fig. 1, which plots mean percentage of repetition responses as a function of trial blocks. Acquisition is rapid, but marked reversals occur in the third block of 10 trials for Groups 30 (alternation groups), and in the fifth block of 10 trials for Groups 70 (repetition groups). Terminal response levels, calculated over Trials 151–200, average 79% for Groups 70, and 33% for Groups 30. These values correspond to optimal response levels of 79% and 67%, respectively. No trial rate effect is apparent for Groups 70, but the slower rate conditions of Groups 30 do somewhat poorer in the middle 100 acquisition trials. There is an immediate change in the transfer trials, in which all groups received the same purely random sequences. However, Groups 30 and 70 show little sign of approaching a common final level, even after the 300 transfer trials given the two 5-sec. groups.

Analyses of acquisition performance on total responses over Trials 1–200,

[3] Between .003 and .004 of the responses were "improper," with *S* pressing too late, pressing neither key, or, most frequently, pressing both keys. Frequency of such improper responses did not appear to depend on trial rate, conditional probability, or trial number, except that Groups 0 and 100 made no improper responses in acquisition. All improper responses were filled in randomly in order to simplify the tabulations.

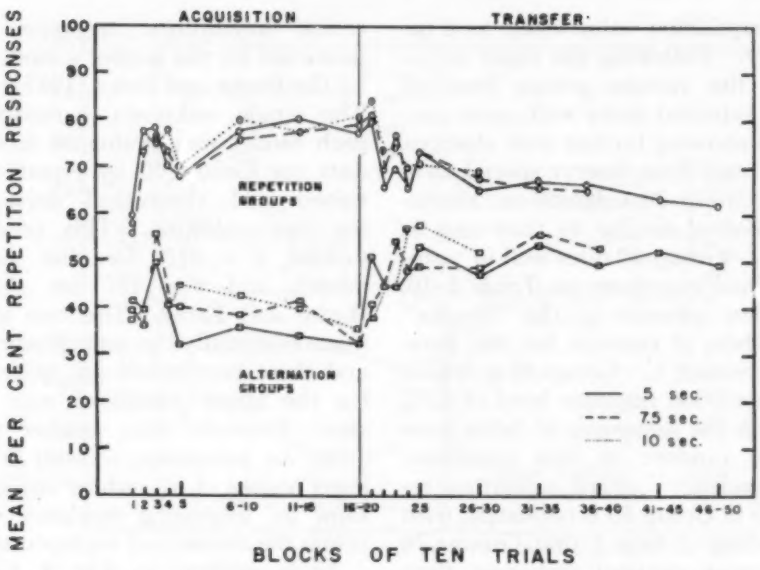


FIG. 1. Mean percentage of repetition responses in acquisition and in transfer as a function of trial blocks, Exp. I.

and over Trials 151–200 were performed for Groups 30 and Groups 70 separately. There were no significant results in the analyses of Groups 70, although the Sequence effect approached the .05 level in the first analysis. For Groups 30, the Sequence effect was significant in both analyses: F (df = 3, 48) = 5.46, and 4.40, for the analyses over Trials 1–200, and Trials 151–200, respectively. No other reliable differences among Groups 30 were found, although the linear component of trial rate approached significance, F (df = 1, 48) = 3.82 in the analysis of the total score over Trials 1–200. The comparison of optimal response frequencies, based on the pooled error term, yielded significant differences between the two conditional probability conditions: F (df = 1, 96) = 97.7 and 51.1, for the tests over Trials 1–200 and Trials 151–200, respectively.

Analyses of transfer performance were made on repetition response frequencies, and included Sequences as an additional factor orthogonal to the two main independent variables. The effect of the conditional probability factor was significant in all analyses. In particular, total repetition responses over Trials 451–500 for the two 5-sec. groups were significantly different, F (df = 1, 32) = 5.15. No other main effect or interaction approached significance.

The first 50 trials of the sequences of reinforcing lights were inspected for peculiarities which might explain the reversals early in acquisition. The reversal for Groups 70 apparently arose entirely from a run of seven consecutive alternations which occurred in one of the sequences. Nothing unusual was seen in the sequences for Groups 30, however. See also below.

Experiment II.—Figure 2 plots the mean percentage of repetition responses as a function of trial blocks

with acquisition value of π_{11} as a parameter. Following the rapid acquisition, the various groups level off in the expected order with some conditions showing further slow changes. Three conditions deserve special mention. Group 30 exhibits an anomalous reversal similar to that seen in Exp. I. Group 40 does well in terms of optimal responses on Trials 1–10, but then ascends to the "chance" level where it remains for the duration of Session 1. Group 50 stabilizes at a repetition response level of 62% although the sequences of lights were entirely random in this condition. This tendency toward repetition responses in Group 50 is consistent with the finding of Exp. I that Groups 70 made more optimal responses than Groups 30. This latter result appears again in Exp. II, and the same trend is shown in the comparisons of Groups 40 and 60, and of Groups 20 and 80.

Table 1 gives the means and *SE*s obtained from total repetition response frequencies over Trials 201–300. This table includes the theoretical asymptotic response levels predicted by the author's model, and by the Burke and Estes (1957) model. The single unknown parameter in each model was evaluated from the data for Group 50 by equating obtained and theoretical asymptotes for this condition. This procedure yielded $c = .618$ for the present model, and $\theta = .472$ for that of Burke and Estes. The two models make essentially the same predictions, and these predictions are quite good for the larger values of π_{11}. It is clear, however, that neither model gives an adequate account at the lower values of π_{11}, where the groups show an increasing tendency to lie below the theoretical asymptote.

TABLE 1

OBTAINED AND PREDICTED MEAN PROPORTION OF REPETITION RESPONSES ON TRIALS 201–300, EXP. II

Group	Obtained	Predicted		s_m
		Eq. 1	Eq. 2	
90	.94	.94	.94	.011
80	.90	.87	.88	.014
70	.84 (.77) (.79) (.80)	.79	.80	.014
60	.75	.71	.71	.020
50	.62	.62	.62	.021
40	.49	.52	.51	.025
30	.35 (.33) (.33) (.35)	.41	.40	.014
20	.16	.29	.28	.020
10	.04	.15	.14	.007

Note.—Entries in parentheses are from Trials 151–200 of Exp. I.

As is evident in Fig. 2, Groups 0 and 100 were shifted to the transfer condition after only 100 acquisition trials. The main results for these two conditions will be given in this paragraph and they will not be considered again unless explicitly mentioned. Mean errors in acquisition were 1.35 and 2.95 for Groups 0 and 100, respectively. The difference is significant as shown by the Mann-Whitney U test, $z = 2.93$. The groups overshoot each other in the fourth block of transfer trials, and Group 0 makes significantly more repetition responses in that block, $F\ (df = 1, 32) = 6.59$ by an ad hoc test. The response rate over Trials 201–500 was quite stable, but these data are omitted in Fig. 2. Group 0 averaged 3.7% fewer repetition responses than Group 50, and Group 100 averaged 2.3% more repetition responses than Group 50 over these last 300 trials. However, the between-groups F for these data was not significant.

In the transfer trials, all conditions show some initial change in response rate except for Groups 40, 50, and

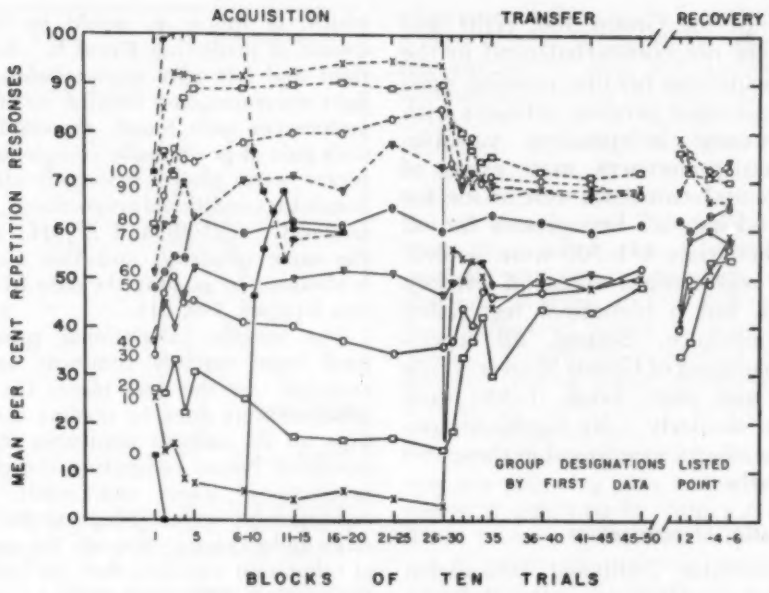


FIG. 2. Mean percentage of repetition responses in acquisition, in transfer, and in recovery, as a function of trial blocks, Exp. II.

60. Group 10 exhibits a sharp increase comparable to Group 0. The terminal transfer levels appear to be undergoing little change. However, these terminal levels are not smoothly related to acquisition treatment: Groups 60–90 all cluster around the 70% level, Groups 10–40 level off at about 50%, and Group 50 falls neatly between.

An analysis of total repetition responses over Trials 401–500 was performed with sequences of reinforcing lights and acquisition treatment as orthogonal factors. For acquisition treatment, F (df = 8, 188) = 16.8, but no other effects were significant. Using the value s_m = 2.64 (for a group of 20 *Ss*) obtained from this analysis in Duncan's (1955) range test, Groups 10–40, Group 50, and Groups 60–90 separated out as three significantly different subsets of conditions, with no further differentiation within these three subsets.

Recovery is evident for most groups in Fig. 2, but dissipates in about 20 trials. Although Session 2 was too short to be more than suggestive, there seems to be a tendency for Groups 60–90 to return to their final level in Session 1, and for Groups 10–40 to rise above their final level in Session 1.

Recovery scores were computed by taking the absolute value of the number of repetition responses on Trials 481–500 from twice the number of repetition responses on Trials 1–10 of Session 2. A positive score thus indicates a recovery of the acquisition response. Groups 0, 50, and 100 were tested together but the analysis yielded no significant results. The mean recovery score, averaged over the remaining eight groups, was significantly different from zero, F (df = 1, 173) = 32.3. Further tests are vitiated by the lack of control over the interval between sessions.

Except for Group 100, right and left were not counterbalanced in the light sequences because position preferences cannot produce artifacts with the present independent variable. Position preferences may exist, of course, and two tests were made for them. First, all key presses for all *S*s over Trials 451–500 were pooled. There was a slight excess of left key presses, but a binominal test failed of significance. Second, all repetition responses of Group 50 over Trials 1–50, and over Trials 1–500 were treated similarly. No significant positional effects were found in these two tests either.

Sequential Dependencies

Considerable additional information may be extracted from the data by considering the dependence of the response on any trial upon the subsequence of stimuli and responses on the trials immediately preceding. The analysis of these sequential dependencies has proved useful in previous work on specific models (Anderson & Grant: 1957, 1958; Estes & Straughan, 1954). Without a specific model, however, the interpretation of the sequential dependencies presents certain pitfalls which have given rise to some confusion in the literature. Accordingly, some comments will be made at appropriate points on limitations in the interpretation of such data.

The two stimulus events will be denoted by 0 and 1. A subsequence of consecutive events will be called a *tuple*. The symbol, $R'(T)$, will denote the proportion of repetition responses conditional on the occurrence of any particular tuple, T. Thus, $R'(100)$ denotes the proportion of repetition responses, given that the stimulus lights on the three preceding trials were 1, followed by 0, followed by 0. Its value was calculated by dividing the number of occurrences of the stimulus tuple, 100, into the number of times that this tuple is followed by a repetition response which, in this case, would be the response of predicting Event 0. Because right and left were equiprobable in all light sequences, and because no position preferences were found, the data from each pair of positionally complementary tuples were pooled before dividing to obtain the conditional proportions. Consequently, $R'(100)$ and $R'(011)$ denote the same quantity, and this quantity is obtained by pooling the data from the two 3-tuples, 100, 011.

The specific tabulational procedure used here requires comment in two respects. In the first place, the compilations were done by pooling the data over all *S*s and all sequences in each condition before computing the desired proportions. There can result, as a consequence, certain biases as discussed more fully below. Second, the method of tabulation was such that, for instance, the tuple, 0000, contributed one instance to $R'(0000)$, two instances to $R'(000)$, three instances to $R'(00)$, and four instances to the repetition response rate, $R'(0)$. Similarly, the alternation tuple, 1010, contributed one instance to $R'(1010)$, two instances to $R'(010)$, three instances to $R'(10)$, and four instances to the repetition response rate, $R'(0)$; (see penultimate sentence of preceding paragraph). This procedure averages the R' values over all possible antecedents of the corresponding tuples, and thus places the R' values for runs, alternations, and other tuples on an equal footing. A somewhat different tabulational procedure has been used by Nicks (1959) in that part of his report dealing with run curves. The present method leads to simpler mathematical expressions relating sequential dependencies and model parameters. However, Nicks' method may prove to be more useful for nonparametric analyses.

It may be worth mentioning that these tabulations, especially those involving response dependencies, are time consuming. Consequently it is advisable to plan the general tabulation procedures and the data recording sheets with care so as to maximize efficiency. The use of the symbols 0 and 1, rather than

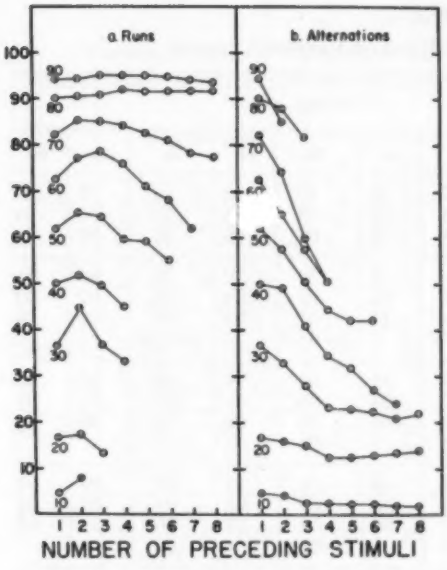


FIG. 3. Mean percentage of repetition responses over Trials 101–300, inclusive: (*a*) response following runs of the designated length; (*b*) response following alternation tuples of the designated length. See text.

R and L, is recommended since they are easier to write and to read.

Responses following runs and alternation tuples.—The proportion of repetition responses following runs and alternations tuples of various lengths were tabulated from the data of Trials 101–300 for each group. The results are given in Fig. 3. Tuples of length greater than eight are omitted, as well as all data points based on fewer than 100 instances in the denominator. Figure 3a plots the data for runs of various lengths. The successive points on each curve give the values of $R'(0)$, $R'(00)$, $R'(000)$, $R'(0000)$, etc. The inverted bowl shape of the five middle curves is the well known Jarvik (1951) negative recency effect, or gambler's fallacy. In Fig. 3b are shown the data for the alternation tuples. The successive points on each curve give the values of $R'(0)$, $R'(10)$, $R'(010)$, $R'(1010)$, etc. As would be expected, the probability of a repetition response decreases as the length of the alternation tuple increases. Somewhat more striking is the fact that in all groups (except Group 90 for which the relevant datum does not appear) the decrement following the third light in an alternation tuple is greater than the decrement following the second light in such a tuple. This result is not an artifact of the tabulational procedure. Indeed, if each point included data only from those alternation tuples having *precisely* the length specified on the abscissa, an even greater difference in the decrements would have been obtained. Finally, by comparing the two panels of Fig. 3, it can be seen than an alternation tuple has a greater overt effect than does a run.

The variation over trial blocks in the response following runs and alternation tuples is exhibited in Tables 2 and 3. These tables list the differences between the R' values for tuples of lengths 2, 3, and 4, and the R' values for tuples of the next shorter length. The values of $R'(0)$ are included in Table 2 in order to permit calculation of the separate R' values entering into the several differences.

A negative difference in Table 2 indicates that the probability of a repetition response decreases as the length of a run increases, i.e., a negative recency effect. The lowest order differences, $R'(00) - R'(0)$, are negative over the first 50 trials, but thereafter are nearly all positive. The two higher order differences also increase over trial blocks. These results suggest that the negative recency effect adapts out over trials. This adaptation seems to be more rapid for conditions with stronger repetition tendencies in the light sequences.

A positive difference in Table 3 indicates that repetition response probability decreases as the length of the alternation tuple increases. Most of the values of $R'(0) - R'(10)$ are negative in the first block of 50 trials. Thereafter they assume fairly stable positive values except perhaps in the transfer trials for Groups 10–30. The values of $R'(10) - R'(010)$ are fairly constant over trial blocks. These latter values are also uniformly larger than the values of $R'(0) - R'(10)$,

TABLE 2

DIFFERENCES OF R' VALUES FOR RUNS OF VARIOUS LENGTHS

Trial Block	R' and R' Diff.	Group								
		10	20	30	40	50	60	70	80	90
1–50	$R'(0)$	13	27	45	45	55	62	72	79	89
	$R'(00)-R'(0)$	01	−04	−03	−09	−04	−08	02	−02	−01
	$R'(000)-R'(00)$			−08	−12	−08	−08	−02	00	01
51–100	$R'(0)$	06	25	41	48	59	70	78	89	91
	$R'(00)-R'(0)$	01	04	01	01	02	04	06	02	02
	$R'(000)-R'(00)$			−06	−13	−07	−02	00	00	00
	$R'(0000)-R'(000)$			(−05)[a]		(−03)[a]		(−02)[a]		
101–200	$R'(0)$	05	17	38	51	61	70	80	90	93
	$R'(00)-R'(0)$	02	00	07	03	06	04	04	01	01
	$R'(000)-R'(00)$			−09	−06	00	00	−01	00	00
	$R'(0000)-R'(000)$			−02		−03		−02		
201–300	$R'(0)$	04	16	35	49	62	75	84	90	94
	$R'(00)-R'(0)$	05	01	09	00	02	05	03	00	01
	$R'(000)-R'(00)$			−07	02	−02	03	00	01	01
	$R'(0000)-R'(000)$			−03		−07		00		
301–400	$R'(0)$	51	39	45	49	62	69	72	74	70
	$R'(00)-R'(0)$	−04	04	08	02	05	06	07	05	06
	$R'(000)-R'(00)$	03	−07	−01	01	−02	04	00	01	03
	$R'(0000)-R'(000)$			−02		−02		−02		
401–500	$R'(0)$	49	46	50	50	60	68	68	71	69
	$R'(00)-R'(0)$	−01	05	07	01	02	05	06	01	07
	$R'(000)-R'(00)$	00	−03	−02	−01	00	05	03	04	03
	$R'(0000)-R'(000)$			−03		01		−01		

Note.—Decimal points omitted.
[a] Calculated from the data of Trials 1–100.

confirming and extending the results of Fig. 3b.

Response following other stimulus tuples.—It is also of interest to examine the sequential dependencies for stimulus tuples other than runs and alternations. The statistic considered is the difference of R' values of tuples which are alike except for the leftmost (trialwise most remote) stimulus event. These differences thus measure the influence of the trialwise most remote event of the tuple on the current response.

The R' values have unequal reliability since the number of instances of a given tuple varies systematically with acquisition treatment. The R' values presented for 3-tuples are based on at least 150 instances in a block of 100 trials. However, six of the 4-tuples of Table 5 had pooled frequencies between 50 and 100 instances. In the transfer trials, of course, all tuples of a given length are equiprobable, and the R' values of tuples of lengths 2, 3, and 4 are based on approximately 1000, 500, and 250 instances, respectively, for a group of 20 *S*s. It should be noted that pooling data across the sequences within each condition introduces a bias to the extent that the number of instances of each tuple varies across sequences. However, the bias would appear to be negligible in the present results. Of the 24 R' values which were also computed from the individual *S*'s data (see below), the average magnitude discrepancy was .005.

The 2-tuple difference, $R'(00)-R'(10)$, may be obtained by adding corresponding values of $R'(00) - R'(0)$ from Table 2 and $R'(0) - R'(10)$ from Table 3.

TABLE 3

DIFFERENCES OF R' VALUES FOR ALTERNATION TUPLES OF VARIOUS LENGTHS

Trial Block	R′ Differences	Group								
		10	20	30	40	50	60	70	80	90
1–50	$R'(0)-R'(10)$	00	−01	−02	−06	−02	−12	05	−03	−01
	$R'(10)-R'(010)$	01	02	06	09	04	25	20		
51–100	$R'(0)-R'(10)$	00	01	00	01	02	06	15	07	10
	$R'(10)-R'(010)$	03	01	08	05	10	14	17		
	$R'(010)-R'(1010)$			(04)[a]		(05)[a]		(11)[a]		
101–200	$R'(0)-R'(10)$	00	00	03	02	06	08	10	04	08
	$R'(10)-R'(010)$	02	02	05	09	07	08	11		
	$R'(010)-R'(1010)$			04		06		05		
201–300	$R'(0)-R'(10)$	01	01	04	00	02	07	06	00	09
	$R'(10)-R'(010)$	01	00	05	07	07	07	18		
	$R'(010)-R'(1010)$			05		06		13		
301–400	$R'(0)-R'(10)$	−04	04	08	02	04	06	08	04	06
	$R'(10)-R'(010)$	01	10	10	10	10	09	13	13	17
	$R'(010)-R'(1010)$			05		07		10		
401–500	$R'(0)-R'(10)$	−02	05	06	00	01	06	06	01	08
	$R'(10)-R'(010)$	01	10	11	12	06	10	11	19	13
	$R'(010)-R'(1010)$			06		02		11		

Note.—Decimal points omitted.
[a] Calculated from data of Trials 1–100.

TABLE 4

DIFFERENCES OF R' VALUES OF PAIRED 3-TUPLES

Trial Block	Group								
	10	20	30	40	50	60	70	80	90
				(*a*) $R'(000)-R'(100)$					
1–50			−11	−19	−17	−17	−08	−02	06
51–100			−08	−21	−16	−05	00	01	03
101–200			−17	−09	−01	−02	−04	01	00
201–300			−09	03	−04	07	00	04	04
301–400	07	−13	−02	01	−03	07	00	03	06
401–500	01	−05	−03	−01	01	09	06	07	06
				(*b*) $R'(110)-R'(010)$					
1–50	12	08	18	20	09	35	30		
51–100	22	03	26	13	22	25	25		
101–200	15	06	19	19	13	13	16		
201–300	08	03	15	15	13	11	23		
301–400	01	18	18	19	18	17	24	24	30
401–500	01	18	21	21	11	19	20	18	24

Note.—Decimal points omitted.

These differences rise from predominantly negative initial values to level off at reasonably stable and definitely positive values of the order of .1. Except for Groups 10 and 20, there seems to be neither any systematic between-groups variation, nor any change in the transfer trials.

The two 3-tuple differences are presented in Table 4 for each group and for successive blocks of trials. The separate R' values in each difference may be readily evaluated by making use of the data of Tables 2 and 3. The values of $R'(000) - R'(100)$ are mostly negative in the early trials, but tend to become positive in the later trials for the groups with higher acquisition values of conditional probability. This trend presumably reflects the adaptation of the negative recency effect previously mentioned. The values of $R'(110) - R'(010)$, although suggestive of some differences between groups and over trial blocks, are rather the more remarkable for the degree of constancy which they exhibit. The average value is of the order of .2 which, being larger than the value of $R'(00) - R'(10)$, re-emphasizes the influence of alternation tuples on S's response.

In order to obtain some indication of the variability of the scores discussed in the preceding two paragraphs, R' values were computed for the individual Ss of Groups 30, 50, and 70, for Trial Blocks 101–200, and 401–500. The resulting analyses are given in Table 5. In particular, the Fs for Mean show that the R' values differences, averaged over the two blocks of trials and over groups, are significantly different from zero.

The four 4-tuple comparisons were made only for Groups 30, 50, and 70; the results are given in Table 6. The values of $R'(0000) - R'(1000)$ are largely negative, again reflecting the negative recency effect. The values of $R'(0010) - R'(1010)$ are all positive and fairly large. Thus an alternation rather than a repetition three trials back in the sequence of lights still exerts a decided effect on the current response. The two remaining differences are essentially zero for Groups 30 and 50, but tend to be positive for Group 70.

The results of this section give ample evidence of the influence of the more remote stimulus events on the current response. In interpreting these results, however, it should be specifically noted that a nonzero difference implies nothing about S's perception of, or memory for these more remote stimuli. Any model which takes account of the trial-to-trial changes in the response probability would predict nonzero differences (Anderson, 1959; Estes & Straughan, 1954).

Response–response dependencies.—Further information may be obtained by considering the dependence of the current response on the preceding response. Unfortunately, these dependencies are more difficult to interpret than are the stimulus dependencies because of a selection effect, and because of certain

TABLE 5

ANALYSIS OF VARIANCE OF R' VALUE DIFFERENCES FOR GROUPS 30, 50, AND 70

Source	*df*	*F* Ratios For		
		$R'(00) - R'(10)$	$R'(000) - R'(100)$	$R'(110) - R'(010)$
Mean	1	51.9*	4.81*	80.1*
Groups	2	1.18	5.42*	1.80
Error (b)	101	(.045)[a]	(.041)[a]	(.069)[a]
Trials	1	2.75	8.56*	0.23
T × G	2	3.40*	2.59	0.66
Error (w)	101	(.020)[a]	(.036)[a]	(.032)[a]

* $P < .05$.
[a] Error mean squares.

TABLE 6

DIFFERENCES OF R' VALUES OF PAIRED 4-TUPLES

Trial Block	Group 30	Group 50	Group 70	Group 30	Group 50	Group 70
	(*a*) $R'(0000)-R'(1000)$			(*b*) $R'(1110)-R'(0110)$		
1–100	08	−10	−07	05	00	04
101–200	−03	−06	−07	−02	01	12
201–300	−06	−12	00	00	03	05
301–400	−04	−04	04	−02	−05	09
401–500	−07	01	−02	01	00	06
	(*c*) $R'(0010)-R'(1010)$			(*d*) $R'(1100)-R'(0100)$		
1–100	10	11	17	07	−03	07
101–200	12	11	06	00	01	04
201–300	16	10	17	01	02	02
301–400	11	15	20	00	−01	03
401–500	11	03	20	−06	−03	−03

Note.—Decimal points omitted.

biases. These two difficulties will be discussed in the appropriate places.

The simple first-order response–response dependency is found by counting the frequency with which the key press on one trial is repeated on the next trial. This statistic was computed for each *S* of Groups 30, 50, and 70, over Trials 401–500, with resulting mean proportions of .538, .533, and .538, respectively. The three values do not differ significantly among themselves, but their mean is greater than the chance value of .500, $F\ (df = 1,\ 101) = 17.6$. This same statistic was also calculated for each of the ten 50-trial blocks for Group 50, but there was very little trial variation. The mean over all 500 trials was .538.

The result of the previous paragraph does not necessarily imply either that *S*s had a set or tendency to repeat the previous key press, or that the previous response has any determining influence on the present response; it could equally well represent the workings of the laws of probability according to the following argument. The probability of a left key press, averaged over trials on which a left key press occurred, will be greater than this same quantity averaged over trials on which a right key press occurred. It may then be expected (at least in the present case; see also below) that the probability of a left key press, averaged over trials following a left key press, will be greater than this same quantity averaged over trials following a right key press. The response will thus tend to repeat itself so that there is no present reason to believe that the obtained value of the response–response dependency represents anything more than this response selection effect. Any mathematical model would, of course, predict the magnitude of this response selection effect and thereby make use of the obtained data as a test of goodness of the model (e.g., Anderson, 1959; Anderson & Grant, 1958). It should be noted that, in general, some response selection effect will always be obtained whenever successive responses are not, for whatever reason, statistically independent. Consequently, the joint dependencies to be considered next are subject to the same difficulties of interpretation.

The dependence of the current response on the joint occurrence of the previous response and the four previous stimuli was also investigated. The notation for these dependencies is similar to that used for the stimulus dependencies. In the argument of each R' value, the symbols preceding the semicolon denote

the previous stimuli in the same way as before, and the symbol following the semicolon denotes the previous response. Thus, $R'(1010; 1)$ denotes the proportion of repetition responses given that the last four stimulus lights were 1,0,1,0, in that order, and that the last response was 1. In each case, the last listed stimulus is the reinforcing stimulus for for the listed response. As before, the data were pooled over key positions so that, for instance, $R'(1010; 1)$ and $R'(0101; 0)$ denote the same quantity. There is thus no restriction involved in presenting the data in such a way that the last stimulus is always 0, and this convention has been adopted. Consequently, the R' values always denote the conditional probability of the response 0 on the next trial.

These R' values were obtained by pooling the data over *S*s and over trials for Trial Blocks 1–300 and 301–500. The comparisons presented in Table 7 are differences between R' values based on the same stimulus 4-tuple, and differing only in the previous response. Nonzero differences correspond, therefore, to predictability conferred by knowledge of the previous response over and above predictability gained by knowledge of the four previous stimuli. Table 7 is so arranged that the last response was correct in the R' value on the left side, incorrect in the R' value on the right side of each difference. In addition, the first four lines are for those cases in which the last two stimuli were alike, and the last four lines are for those cases in which the last two stimuli were unlike.

The problem of bias must be considered before discussing the results. Since the R' values are proportions conditional on the previous response, they involve a denominator which is subject to variation. This random variation introduces a bias which is approximately inversely proportional to sample size (Cramer, 1946, Sec. 27.7). An R' value computed for an individual *S* will thus be biased because of the limited number of responses made by that *S*. If *S*s were identical in the sense of having the same learning rates, this bias could (and should) be reduced by pooling the data of the group in order to compute a single ratio. However, if *S*s are not identical, this pooling procedure will itself introduce a bias because of the limited number of *S*s. This bias will be found even when there is sufficient data for each *S* to render negligible the bias in the individual ratios. In fact, the bias in the individual scores reported below is negligible; however, the entries of Table 7, which are based on pooled group data, do involve appreciable bias.

TABLE 7

DIFFERENCES OF PAIRED R' VALUES FOR STIMULUS-RESPONSE DEPENDENCIES

R' Differences	Group 30		Group 50		Group 70	
	Trials		Trials		Trials	
	1–300	301–500	1–300	301–500	1–300	301–500
$R'(0000; 0) - R'(0000; 1)$	51	37	37	34	34	35
$R'(1000; 0) - R'(1000; 1)$	21	26	20	15	22	22
$R'(0100; 0) - R'(0100; 1)$	04	09	13	21	15	12
$R'(1100; 0) - R'(1100; 1)$	07	10	10	21	11	04
$R'(1110; 0) - R'(1110; 1)$	−02	−02	−03	02	04	00
$R'(0110; 0) - R'(0110; 1)$	−05	04	−07	−07	−06	−01
$R'(0010; 0) - R'(0010; 1)$	−04	00	−06	−06	05	18
$R'(1010; 0) - R'(1010; 1)$	−07	01	−09	−04	−02	−11

Note.—Decimal points omitted.

This observed bias may, therefore, be taken as reflecting the existence of individual differences.

It may be noted, parenthetically, that any model which took into account the trial-to-trial changes in response probability would yield theoretical expressions (possibly of a Monte Carlo variety) for the *unconditional* proportions. The use of such data would avoid the problem of bias arising from the use of ratios with a random variable denominator. Since the two models considered here have already been seen to be inadequate, this procedure has not been followed. Without a model, however, the unconditional proportions are uninterpretable. Consequently, the nonparametric procedure adopted here, despite its defects, seems necesary to obtain meaning from the results.

The results for Group 50 will be considered first since they are the most reliable. The entries in the first four lines of Table 7 are positive and large. Positivity would be expected on the basis of the response selection effects discussed above since the previous key press in the R' values on the left of each difference is a 0, and since the repetition response is a 0 for both R' values of each difference. However, the Table 7 entries seem considerably larger than one would expect if the response selection effect alone were operating. The most appropriate measure of the size of the response selection effect is obtained by doubling the difference between the simple first-order response–response dependency and the chance level. The previous data of this section thus yield the value of .076 against which to compare the Table 7 entries. This comparison value may actually be too large since taking into account the four previous stimuli, as is being done here, would be expected to reduce the amount of information conferred by knowledge of the previous response.

In order to test this comparison, the bias in the entries of Table 7 must be allowed for. Two values of $R'(000; 0) - R'(000; 1)$ were calculated for each S of Group 50, each value based on the data of alternate 50 trial blocks. Since the bias is inversely proportional to the amount of data, comparing the mean of these two values with the corresponding value based on all the data for that S yields a measure of the bias in the individual scores arising from the use of the ratio. This bias was negligible. It may be noted, incidentally, that the correlation between the two scores was .55. The mean of the individual values of $R'(000; 0) - R'(000; 1)$ was .24 which is significantly larger than the comparison figure of .076, $F(df = 1, 39) = 50.83$. Also, the .24 value is less than the weighted mean of .27 obtained from the Group 50 entries in the first two lines of Table 7. It is thus seen that these particular entries are biased away from zero.

The finding that the response selection effect is insufficient to account for the Group 50 entries in the upper half of Table 7 should be interpreted cautiously. It may be necessary to allow for a determining effect of the previous response on the present response as is done in the experimenter–subject controlled models of Bush and Mosteller (1955). A second possibility is that the assumption of path independence must be abandoned as in the models proposed by Sternberg (1959b).

The Group 50 entries in the lower half of Table 7 are predominantly negative. Statistical significance was investigated by computing individual values of $R'(10; 0) - R'(10; 1)$ according to the procedure used for the upper half of the table. The split-half analysis yielded a correlation of .67, and showed also that the bias in the individual scores was negligible. The mean of the individual scores was $-.036$, with a standard error of .030. The tabular entries are again biased away from zero, since the corresponding mean obtained from Table 7 was $-.053$. Although the Group 50 entries in the lower half of Table 7 are not significantly less than zero, they deserve comment since the response selection effect noted above would, if taken at face value, imply that these entries should be somewhat greater than

zero. In accord with the original model, and with the preceding analysis of stimulus dependencies, it is assumed that *S*s have an alternation response (predicting next that light which did not flash last) in their repertoire. Now, in the R' value on the left of each difference, an alternation response (a 0) has just been made (and rewarded). But in the R' value on the right side of each difference, a repetition response has just been made (and punished). Hence it would be expected on purely probabilistic grounds that an alternation response (a 1) would be more likely on the next trial in the first case than in the second. This means that a repetition response on the next trial will be less likely in the first case than in the second, in agreement with the data. By thus interpreting the overt key presses as alternation and repetition responses, it is seen that the simplest explanation of the negative values is in terms of a response selection effect. Conversely, the fact that the entries are not positive gives added support to the existence of the alternation response. The further interpretation of the results as indicative of differential effects of reward and nonreward would be premature.

The data for Groups 30 and 70 need little discussion. The entries in the upper half of Table 7 present essentially the same picture as Group 50 did. The data for the lower half of the table do not show such good agreement. This might perhaps be expected for Group 70 since the alternation response presumably does not develop as strongly there. For Group 30, individual values of $R'(10;0) - R'(10;1)$ over Trials 1–300 were computed using the split-half technique as before. The correlation of the two scores was .69, and again the bias in the individual scores was negligible. The mean individual score was −.061 with a standard error of .031. While the mean thus falls slightly short of being significantly less than zero, it reinforces the corresponding analysis of the Group 50 data. It should be noted, however, that the values over Trials 301–500 for Group 30 are not negative, a result which is contrary to the above explanation.

Miscellaneous Results

This section gives two results from the combined data of Exp. I and II.

It has been observed that Groups 30 in both experiments show a hump at the third block of 10 acquisition trials, first increasing and then decreasing in frequency of repetition responses. Tests were made on the difference of the total scores on Trials 1–10, and Trials 21–30. The mean difference was significantly greater than zero both for Exp. I, $F\ (df = 1,\ 48) = 18.81$, and for Exp. II, $F\ (df = 1,\ 24) = 12.46$. Moreover, there were no complications from other factors: in Exp. I, neither trial rate, stimulus sequences, nor their interaction were significant; in Exp. II, where eight different sequences were used, the F for Sequences was again close to unity. Group 40 lay significantly below the 50% level on the first block of 10 trials, which is consistent with the behavior of Group 30.

These results suggest the presence of an initial tendency toward alternation responses. In order to obtain a purer measure of initial tendencies, the data from the first two key presses are given in Table 8. Here the first stimulus light follows the first key press and precedes the second key press. Table 8 includes the present data, that from Exp. 3 of

TABLE 8

JOINT DISTRIBUTION OF FIRST TWO KEY PRESSES AND FIRST REINFORCING STIMULUS

First Two Key Presses	First Reinforcing Stimulus		Total
	Right Light	Left Light	
Right-Right	119	114	233
Right-Left	99	56	155
Left-Right	40	44	84
Left-Left	85	47	132
Total	343	261	604

Note.—The reinforcing stimulus follows the first key press and precedes the second key press.

Anderson and Grant (1957), and that from a pilot study, to yield a total N of 604. In all cases, *Ss* had three previous practice trials in which both lights or neither light could come on, but not either light singly. There is a significant preference for the right key on Trial 1, but not on Trial 2. The probability that the second press is an alternation response is .57, significantly above chance by a binomial test. It will be noted that the data could also be interpreted in terms of the "win–?, lose–stay" strategy of Goodnow and Pettigrew (1955). A somewhat different treatment of these data has been made by Sternberg (1959a) in terms of a model embodying a perseverative tendency.

Discussion

This discussion will first compare the present results with those of other experiments, and then consider the mathematical model approach to probabilistic learning situations of the kind investigated here.

The most relevant comparison data are those of Engler (1958). Four conditions in her Exp. II had acquisition values of π_{11} identical with those of the present Groups 10, 50, 70, and 90. The numerical values of the terminal acquisition levels are in fairly good agreement except that, for Group 10, the present asymptote of .04 is considerably lower than the .11 figure obtained by Engler. Her Group G, with $\pi_{11} = .25$, also terminates somewhat higher than might be expected by interpolation in Table 1. The source of this discrepancy is not clear. However, it is somewhat disturbing in view of the generally good agreement among different experiments on asymptotic values for the frequency variable.

It is worth mentioning that Groups 0 and 100, which received continuous reinforcement in acquisition, were little different from Group 50 in the transfer trials, but had significantly different transfer asymptotes than the other alternation and repetition groups. This result is not inconsistent with the common finding of greater resistance to extinction after partial reinforcement if the transfer condition is considered as analogous to extinction. It is possible, of course, that Groups 0 and 100 would have reached different transfer asymptotes had they been given more than 100 acquisition trials.

The finding that Group 0 learned faster than Group 100 is the reverse of the result of Goodnow and Pettigrew (1956). A plausible explanation lies in the difference in the reinforcing stimulus situation. In their "two-armed bandit" task, the fact that the other response would have been correct on trials when *S* received no chip was established by instruction, but no special signal to this effect was given during the experiment as was done in the procedure here.

The results from the alternation tuples do not seem to be entirely in agreement with corresponding results of Nicks (1959). In particular, $R'(0010) - R'(1010)$ has the value of about .2 for Groups 30, 50, and 70, (Table 6). Nicks' Table 2, however, shows values of the same statistic of .02 and .04. The tabulation procedures were the same in both cases, and the reliabilities of the data are such as to make the discrepancy significant. Nicks reports these data only for his 67:33 frequency condition so that the source of the discrepancy may stem from this difference in the experimental situations. In any event, the present data show that Nicks' emphasis on runs as the only important stimulus tuples is not justified in general.

The final comparison with existing data is restricted to Group 50. For this condition, a statistic obtained by doubling the difference between the repetition response level and the chance level is the same as a sequential statistic computed by Estes and Straughan (1954). This quantity has here an asymptotic value of about .24 which agrees with the average value of .24 obtained from their three frequency conditions. In the report of Anderson and Grant (1957), however, these values (there denoted by a) tended to be somewhat larger. In all three of these ex-

periments it was found that the a values were considerably lower in the early trials.

The predictions made by the present model, as well as by the revised model of Burke and Estes (1957), disagree with the experimental results in a number of respects. In the first place, the acquisition asymptotes at the lower π_{11} values are considerably below the theoretical values. Second, the analyses of the sequential dependencies have exhibited several deviations: the negative recency effect, which is commonly reported, was again found here; the alternation tuple data showed that the third preceding light in the stimulus tuple, *RLR*, had a greater effect than the second preceding light in the stimulus tuple, *RL*, a result which is contrary to both models; in addition, the analysis of the response dependencies suggested that knowledge of the previous response yielded more information about the present response than would be expected theoretically, at least for models similar to those considered here. Finally, the fact that the final levels in the transfer condition (in which all groups received the same treatment) were different, and showed no sign of converging, is definitely contrary to the models.

Although these discrepancies have been obtained in the conditional probability situation, it is evident that they are in large part applicable to models for the frequency situation as well. In particular, the $\pi_{11} = .5$ condition is common to both the frequency and the conditional probability situations. Thus, although the conditioning process assumed in current mathematical models for the two-choice task may well occur, it would appear that the models neglect much of the underlying processes. Furthermore, since the mean learning curve comprises an average of the sequential increments and decrements in response probability, the agreement between obtained and predicted "matching solution" behavior must be considered fortuitous.

A necessary condition for constructing an adequate model is brought out by the sequential dependency analyses. The negative recency effect and the alternation tuple results imply specifically that it will be necessary to take into account the reinforcing stimulus on more than the one preceding trial which was considered in the two models discussed here. The situation could then be considered as discrimination learning, with memory traces of the last several events serving as discriminative stimuli in a system of S-R associations. It would probably be desirable to allow an analogous trace representation of preceding responses in the model, although the existing evidence is not unambiguous on the necessity for this.

The formulation of the preceding paragraph is consistent with the claim of Hake and Hyman (1953) that *S*s perceive and respond to specific sequences of events preceding each trial, and is perhaps not opposed to the notion of "strategies" used by Goodnow and Pettigrew (1955). However, it is apparent that the elucidation of the effective discriminative stimuli will not be as simple as has been thought. The specification of these stimuli will probably rely heavily on sequential dependency data, but these, as has been seen, include a number of automatic effects not directly indicative of underlying psychological processes. Indeed, it would appear that the dissection of these data cannot be carried very far without an exact model. Estes' (1950) conceptualization of the stimulus, together with the treatment of probabilistic discrimination learning given by Burke and Estes (1957), affords a possible method of attack. Thus, for instance, shorter tuples or tuples of greater perceptual impact would correspond to larger subsets of stimulus elements or to subsets of stimulus elements with greater sampling probabilities. The stimulus element approach would also be well suited to quantifying the similarity of the total stimulus trace on successive trials. Sternberg's (1959b) path-dependent models with many-trial dependencies represent a second possible approach to the problem.

It was expected that the sequential

dependency analyses would give the most incisive information about the behavior since they most directly reflect the trial-to-trial increments and decrements in the variables underlying the overt response. Nevertheless, the portion of these data which was immediately useful was not overly large. As has been seen, the nonparametric interpretation of such data is difficult when the behavior is as complicated as is the case here. The sequential dependencies are most useful in testing or evaluating parameters of particular models since any quantitative model must predict their magnitude. A general mathematical treatment of sequential dependencies for linear learning models has been given by Anderson (1959). It can be shown, for example, that the original Estes and Straughan (1954) model implies that the entries within each of Tables 5 and 6 should be constant and equal to $\theta(1-\theta)^2$, and $\theta(1-\theta)^3$, respectively. Corresponding expressions for the two models discussed here were not worked out since they failed on other grounds. However, the cited expressions point up the usefulness of the dependency data. It is hoped that the present results will be of use in testing future models.

The behavior in the transfer condition raises a serious problem for any theoretical treatment of the two-choice probabilistic situation. It will be recalled that in the transfer condition, in which all *S*s received a common treatment, the various groups showed no sign of converging to a common level of repetition responses even after hundreds of trials. The data also indicated that this transfer effect was not smoothly related to acquisition treatment. From a qualitative point of view, it is apparent that the acquisition treatment has induced some permanent or semipermanent change in the *S*s. However, the results of the sequential dependency analyses have shown that at least some of the various stimulus tuples must be considered separately. Consequently, since the induced change could be associated with any or all of these tuples, it is not possible to get a clear specification of its locus on the basis of the existing data.

The same problem holds for the mathematical model approach, but in an even more acute way since quantitative results are desired. The trace stimulus formulation would be feasible if the traces are short range and do not extend back more than, say, a half dozen trials. However, the transfer behavior is a long-range effect spanning several hundred trials. It is clearly impracticable to allow directly for such long term traces in any model. The same difficulty would apply to Sternberg's (1959b) path-dependent models.

It is possible, of course, that the long range effects arise from systematic changes (e.g., Restle, 1955) in the populations of stimulus elements which give rise to the short range traces. If this is correct, the suggested development would still be practicable although it would require the incorporation of the laws of such stimulus change. There is some evidence which may very gingerly be interpreted as suggesting that such laws of stimulus change may not be too complicated. The fact that the between-groups differences in the higher order dependencies were not more marked indicates that the changes in the stimulus element populations may possibly be largely confined to the stimulus 2-tuples. However, such a simple resolution of the problem of long range effects would still be insufficient to salvage the two models discussed here.

It seems not unreasonable to suggest that the most promising present line of attack on the two-choice situation lies in an extensive empirical investigation of the behavior under a variety of transfer conditions. This would include the study of repetition responses in situations in which the light frequencies are the independent variables. The present data, and those of Engler (1958) and of Goodnow and Pettigrew (1955) who also used a transfer paradigm, cover only a small part of the picture. Until more knowledge of transfer behavior is available, spot checking of specific mod-

els is liable to be little better than shooting in the dark.

In future work, it might well be more efficient to procure large quantities of data for a few *S*s so that reliable individual values of the sequential dependencies can be obtained. It is apparent from the present results that a number of parameters will be needed to account for the behavior, whether or not this is done in a formal mathematical way. Averaging data over *S*s in order to increase reliability is only too likely to simultaneously confound the descriptive statistics with individual differences (Anderson, 1959). This increases the difficulty of interpreting the statistics as well as any model parameters derived from them.

Summary

Two experiments on two-choice probability learning were reported which varied the first-order conditional probability in the stimulus event sequences. This variable is the probability that an event occurs on one trial, given that it occurred the previous trial. The dependent variable was predictions of that same event which had occurred the previous trial.

Experiment I used a 3×2 design with three trial rates and two values of conditional probability in the 200 acquisition trials. Acquisition was followed by transfer to completely random sequences. The conditional probability groups reached significantly different acquisition asymptotes, and stable, but significantly different, response rates in transfer.

In Exp. II, 11 groups were given acquisition training with conditional probabilities ranging from 0 to 1. Excepting the two extreme groups, 300 acquisition trials, 200 transfer trials, and 60 recovery trials were given. Acquisition asymptote was an orderly function of conditional probability. Stable but unequal response rates in transfer were obtained, as in Exp. I. Spontaneous recovery was also obtained.

Detailed analyses of sequential dependencies in the stimulus-response sequences found, in addition to the "gambler's fallacy," strong effects of strings of consecutive alternations in the stimulus sequences. Some evidence for a determining effect of previous responses was obtained.

Two models based on statistical learning theory predicted asymptotic acquisition response levels quite well for the higher values of conditional probability, but did poorly at the very low values. The models also failed to account for the nonconvergence in transfer, and for much of the sequential dependency data. It was concluded that agreement between obtained and predicted "matching solution" behavior is fortuitous.

References

Anderson, N. H. Effect of first-order conditional probability in a two-choice learning situation. Unpublished doctoral dissertation, Univer. of Wisconsin, 1956.

Anderson, N. H. An analysis of sequential dependencies. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959.

Anderson, N. H., & Grant, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, **54**, 305–317.

Anderson, N. H., & Grant, D. A. Correction and reanalysis. *J. exp. Psychol.*, 1958, **56**, 453–454.

Burke, C. J., & Estes, W. K. A component model for stimulus variables in discrimination learning. *Psychometrika*, 1957, **22**, 133–145.

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Cramer, H. *Mathematical methods of statistics.* Princeton: Princeton Univer. Press, 1946.

Duncan, D. B. Multiple range and multiple *F* tests. *Biometrics*, 1955, **2**, 1–42.

Engler, J. Marginal and conditional stimulus and response probabilities in verbal conditioning. *J. exp. Psychol.*, 1958, **55**, 303–317.

Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

Goodnow, J. J., & Pettigrew, T. F. Effect of prior pattern of experience upon strategies and learning sets. *J. exp. Psychol.*, 1955, **49**, 381–389.

Goodnow, J. J., & Pettigrew, T. F. Some sources of difficulty in solving simple

problems. *J. exp. Psychol.*, 1956, **51**, 385–392.

GRANT, D. A., HAKE, H. W., & HORNSETH, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.

HAKE, H. W., & HYMAN, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45**, 64–74.

JARVIK, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, **41**, 291–297.

NICKS, D. C. Prediction of sequential two-choice decisions from event runs. *J. exp. Psychol.*, 1959, **57**, 105–114.

RESTLE, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62**, 11–19.

STERNBERG, S. H. An analysis of four models for human learning. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press. (a)

STERNBERG, S. H. A path-dependent linear model. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. (b)

WILK, M. B., & KEMPTHORNE, O. Fixed, mixed, and random models. *J. Amer. Statist. Ass.*, 1955, **50**, 1144–1167.

(Received May 4, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

RELATION OF SERIAL POSITION ERRORS TO DOUBLET AND SPLIT-DOUBLET LOCATION IN VERBAL MAZE PATTERN [1]

RONALD L. ERNST, DONALD R. HOFFELD, SIDNEY SEIDENSTEIN, AND W. J. BROGDEN

University of Wisconsin

Thompson, Voss, and Brogden (1957) investigated the effect of doublet location within the verbal maze pattern on its acquisition and the form of the serial position error curves. The 16-unit 4-choice pattern of Brogden and Schmidt (1954) was modified by locating doublets (two successive like responses) at Units 2 and 3, 5 and 6, 8 and 9, 11 and 12, and 14 and 15, respectively, for the five experimental groups. The serial position curves of the difference in errors between each experimental group and the control group show a doublet effect of a significant increase in errors for the second item. Analysis of first-order contingent responses on the first trial shows that *S*s tend not to repeat a response correct for the immediately preceding position and suggests that this kind of guessing behavior during the discovery phase of acquisition accounts for the doublet effect.

The method of producing a doublet for the experimental group patterns generated split-doublets (two identical response items with a different response item between them), of which there were none in the control group pattern. The split-doublets produced a significant increase in error at the position of the second like item (third position of split-doublet sequence). The unintended confounding of doublets with split-doublets in the experimental maze patterns may be responsible for the results. Of the two experiments to be reported[2], one is an exact repetition of the one performed by Thompson, Voss, and Brogden (1957) with elimination of split-doublet sequences from the patterns of the experimental groups. The other experiment was designed to study the effect of location of split-doublets.

PROCEDURE

Experiment I.—The 16-unit maze with the four alternate choices, 10, 20, 30, and 40, each occurring four times in the pattern, was learned by the control group. The five patterns acquired by the experimental groups contained a single doublet at Maze Units 2 and 3, 5 and 6, 8 and 9, 11 and 12, 14 and 15, respectively, and each represented a limited modification of the standard maze to provide a doublet without split-doublets or other noticeable sequences elsewhere in the maze. Each maze was permuted to obtain four subpatterns and thus provides for each number to be represented once at each of the 16 positions.

Of 24 *S*s in each group, six learned one of the four subpatterns. Half of the *S*s in each condition were treated by one of two *E*s. The *S*s were students in elementary psychology at the University of Wisconsin who volunteered to serve. Assignment of *S*s to conditions was random within each replication.

The instructions of Thompson, Voss, and Brogden (1957) were read to *S*. Responses of 10, 20, 30, or 40 were made by *S* with *E*

[1] This research was supported in part by grants from the National Science Foundation and the Research Committee of the Graduate School from funds provided by the Wisconsin Alumni Research Foundation.

[2] Hoffeld and Seidenstein collaborated with the senior author on Exp. I and Ernst on Exp. II.

signaling the correct response by saying "Forward." This procedure continued until *S* completed one errorless trial. Each response at each unit was recorded by *E* for each trial. A stop watch was used to measure the time of each trial and the 1-min. rest between trials. Other conditions were comparable to those of Thompson, Voss, and Brogden (1957).

Experiment II.—The four patterns acquired by the experimental groups contained a single split-doublet at Maze Units 2–4, 5–7, 10–12, and 13–15, respectively, and each represents a limited modification of the standard maze to provide a split-doublet without doublets or other noticeable sequences elsewhere in the maze. Except that a single *E* treated all *S*s (24 in each group, 6 in each subgroup), the procedure for Exp. II was identical to that for Exp. I.

Results

Experiment I.—Analyses of variance of the data in terms of time and trials to the criterion trial showed no significant effect for maze pattern, maze pattern subgroups, *E*, or the interaction of these variables. Since the assumption of homogeneity of variance was found not tenable for the error data, the 1% level was used to determine statistical significance for all analyses of these data. No source from the between-*S*s variation was significant. For the within-*S*s variance, the significance of serial position establishes the reliability of the serial position error curves. Of the significant interactions, only that for serial position by maze pattern is meaningful in terms of the experimental design. Graphic plots of the other interactions (serial position by submaze pattern and serial position by maze pattern by submaze pattern) show differences that appear to be only of random character.

Since there are no significant differences between groups in terms of total errors to the criterion, and since the primary results are the variation in form of the serial position error curve as a function of maze pattern, the lead of McCrary and Hunter (1953) was followed by converting the error data at each serial position to percentage of total error. The number of errors over all trials to the criterion was converted to percentage of total error over all trials and all positions for each of the 16 serial positions for each *S*. The group mean of the percentage scores for *S*s of each experimental group was computed for each serial position and the plot of these values is presented in Fig. 1 by the solid line curves. In order to increase the stability of the standard serial position error curve, data for the control group of Thompson, Voss, and Brogden (1957) were converted to percentages and combined with the data for the present control group. Analysis of variance of the data for the two control groups shows no significant differences between them. Comparison of the serial position curve for all errors in the combined control group (Fig. 1) with the comparable serial position curves for the experimental groups shows substantial differences in the form of the curves. These differences are presented more clearly in the solid line curves in the left half of Fig. 2. The data for these curves were obtained by taking the difference in percentage of error for each *S* and the mean percentage error for the combined control group at each of the 16 serial positions and obtaining the mean of these differences for each of the five experimental groups. For each group there is an increase in error for the second position of the doublet. The difference data were subjected to an analysis of variance resulting in a significant interaction of serial position by maze pattern. This establishes significant differences in the form of the differ-

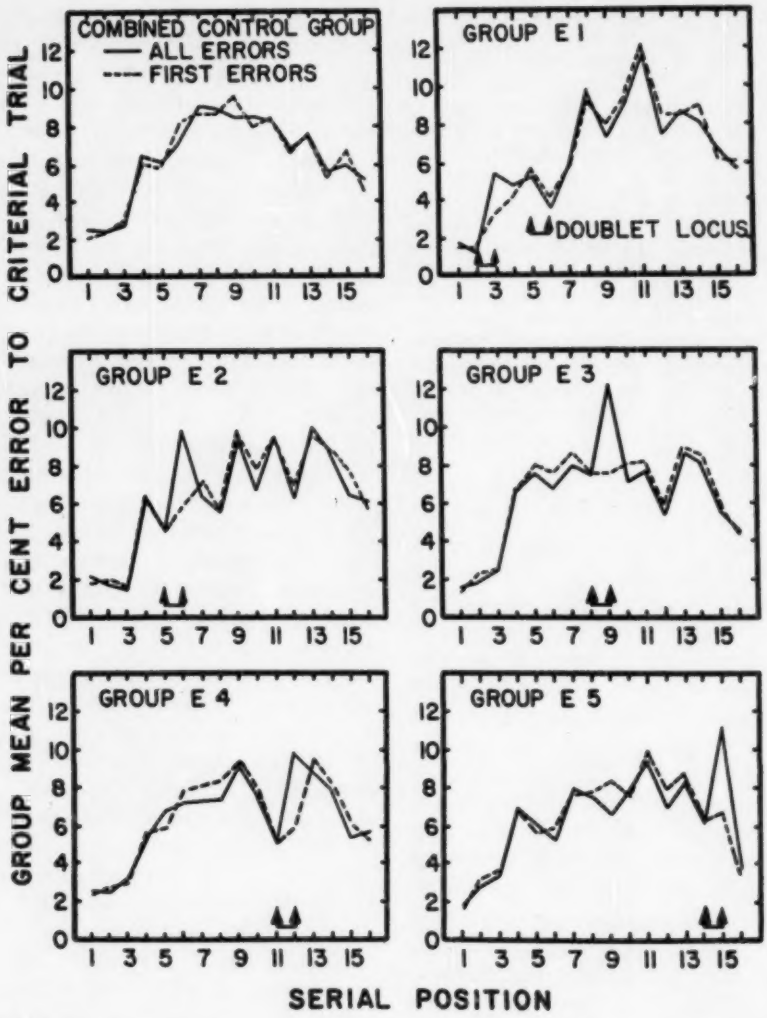


FIG. 1. Serial position curves in terms of percentage of total errors and percentage of first errors as a function of doublet locus.

ence curves between the five experimental groups. The differences in the form of the curve involve the effect of doublet location and other differences attributable to variation in maze pattern required for construction of doublets at appropriate positions for the different experimental groups. The results shown by these curves are comparable to the results of Thompson, Voss, and Brogden (1957) for doublet location and are not confounded by the split-doublet effect that occurred in the earlier study.

Thompson, Voss, and Brogden (1957) demonstrated that frequency of first-order contingent response was reduced for any given number if it was followed by the same number.

The magnitude of this effect was approximately the same for the numbers 20, 30, and 40, but was substantially less for the number 10 (more responses of 10 when the number previously correct was 10). Analysis of first-order contingent responses for *Ss* of the current study produced similar results. Serial position error curves based on the first errors per

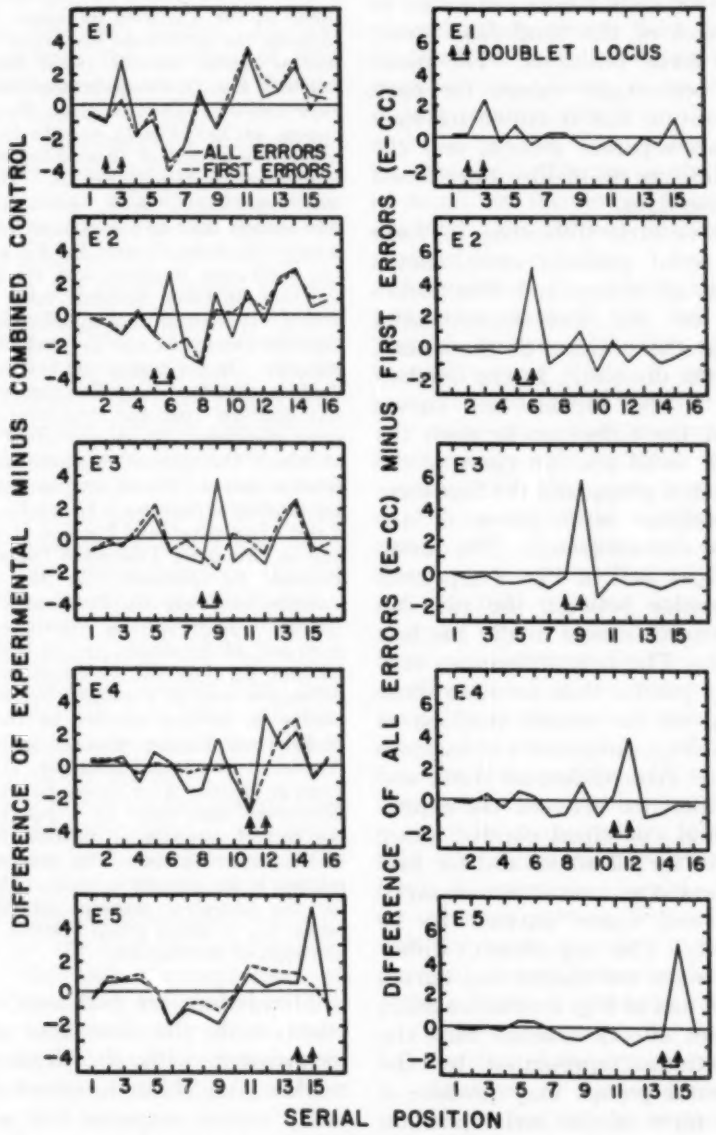


FIG. 2. Curves of differences between serial position curves for experimental minus control groups for percentage of total and first errors (left side) and curves of the differences between these difference curves for each experimental group (right side).

trial should show the doublet effect and provide more information about the relation between the doublet effect and the tendency not to repeat a response previously correct. The totals of the first errors for each position for each *S* were converted to percentages of the total first errors over all serial positions. The mean of the percentage values for each serial position was obtained for each of the six separate groups, and the broken line curves of Fig. 1 are based upon these data.

There is little difference in form of the serial position error curves based on all errors and first errors except for the five experimental groups at the position of the second item of the doublet. In the left half of Fig. 2 the broken line curves represent the difference between the first-error serial position curve of the experimental group and the first-error serial position error curve of the combined control group. The curves of the right half of Fig. 2 represent the difference between the two difference curves shown in the left half of Fig. 2. The only differences substantially greater than zero are those occurring at the second position of the doublet. Analyses of variance of the first-error difference data, and of the difference between the experimental and combined control group differences for all errors and for first errors show the interaction of serial position and maze pattern to be significant. The significant differences between the dotted line curves of the left half of Fig. 2 (which exhibit no doublet effect) indicate that the maze patterns represented by the experimental groups may produce a different form of the serial position error curve at other loci than those for the doublet. The significant differences between the curves of the right half of Fig. 2 apply only to the doublet effect and indicate that this effect is a direct function of doublet location for all-errors measure of the serial position error curves.

The *t* test was used to establish fiducial limits at the 1% level for mean differences between the percentage total error of experimental minus control (solid line curves, left half, Fig. 2), the percentage first error of experimental minus control (broken line curves, left half, Fig. 2), and the first of these minus the second of these differences (solid line curves, right half, Fig. 2). These limits are, respectively, ±2.8, ±2.8, and ±2.2. The second item of the doublet exceeds the positive limit for Groups E_3, E_4, and E_5 for the total-error measure, and for all groups for the difference between total and first errors. It is within .2 of equaling the positive limit for Groups E_1 and E_2 for the total-error measure. In the case of the first item of the doublet, that for Group E_4 exceeds the negative limit for both the total-error and first-error measures. There are other positions at which the various measures exceed the fiducial limits. These are as follows: for the total-error measure in Group E_1, Positions 6 and 7 (negative), and Position 11 (positive), and in Group E_2, Position 8 (negative) and Position 14 (positive); for the first-error measure in Group E_1, Positions 11 and 14 (positive), in Group E_2, Positions 8 (negative) and 14 (positive), and in both Groups E_3 and E_4, Position 14 (positive). These differences may be a complex function of the change in pattern relative to the pattern of the control group required to locate the doublet at the desired positions. The differences at Position 8 for Group E_2 are the only differences that occur at a position where the correct response is different from that of the control group. The only consistent relation is the increase in error at Position 14 for the first-error measure for all groups except E_5, in which group Position 14 is the first item of the doublet.

The failure to find any doublet effect with the first-error data is inconsistent with the demonstrated tendency of *S*s not to repeat a previously correct response and with the hypothesis that the doublet effect is primarily a function of this kind of guessing behavior during the dis-

covery phase of acquisition of the maze. The doublet effect may be present in the first-error data during the discovery phase, but not show up in the data for all trials, either because of its small magnitude or because of a negative doublet effect after the discovery phase. To test this hypothesis the number of first errors for the first quarter and the last three-quarters of trials to the criterion was tabulated for each *S* for the doublet positions for the experimental groups. Similar data were tabulated for the control group at the serial positions for which doublets occurred in the experimental groups. The frequency of first error at the first doublet item was subtracted from that for the second item. From this difference was subtracted the mean difference for the appropriate serial positions for *S*s of the combined control group. This mean over all experimental groups is .7 for the first one-quarter of trials and −.5 for the last three-quarters of the trials. The results of one-tailed *t* tests show the mean difference for the first one-quarter of the trials to be significantly greater than zero and that for the last three-quarters of the trials to be significantly less than zero at the 1% level.

The total errors for the first quarter and last three-quarters of the trials were treated in the same way as the first-error data. The group mean difference over all experimental groups was 4.9 for the first one-quarter and 2.5 for the last three-quarters of trials. One-tailed *t* tests show both means to be significantly greater than zero at the 1% level, and the mean of 4.9 to be significantly greater than the mean of 2.5.

A further analysis of the doublet effect was attempted by constructing serial position curves from the trial data. The number of trials to and including the last error for each position was tabulated for each *S* and was converted to percentage of trials to the criterion. Examination of curves for group means and for mean differences between each experimental and control group shows them to be of similar form to the curves constructed from errors. There is no evidence of any doublet effect for Groups E_1 and E_5, but for Groups E_2, E_3, and E_4 there is a decrease in the number of trials for the second item of the doublet. No further analyses were made of the trial data.

Experiment II.—The trials, time, and error data of the four experimental groups and the combined control group (Exp. I) were subjected to analysis of variance. Since in each case the assumption of homogeneity of variance was not tenable, the 1% level was used to determine statistical significance for these and all subsequent analyses. There were no significant differences for the trial or time data. In the analysis of the error data, no source from the between-*S*s variation was significant. For the within-*S*s variance, the significance of serial position establishes the reliability of the serial position error curves. Of the significant interactions, only that for serial position by maze pattern is meaningful. Graphic plots of the other significant interactions (serial position by submaze pattern, and serial position by maze pattern by submaze pattern) show differences that appear to be of random character only.

Further treatment of these data follows that of the error data in Exp. I. Analyses of variance were completed on percentage of total error, difference in percentage of total error between the experimental groups and the control group, percentage of first

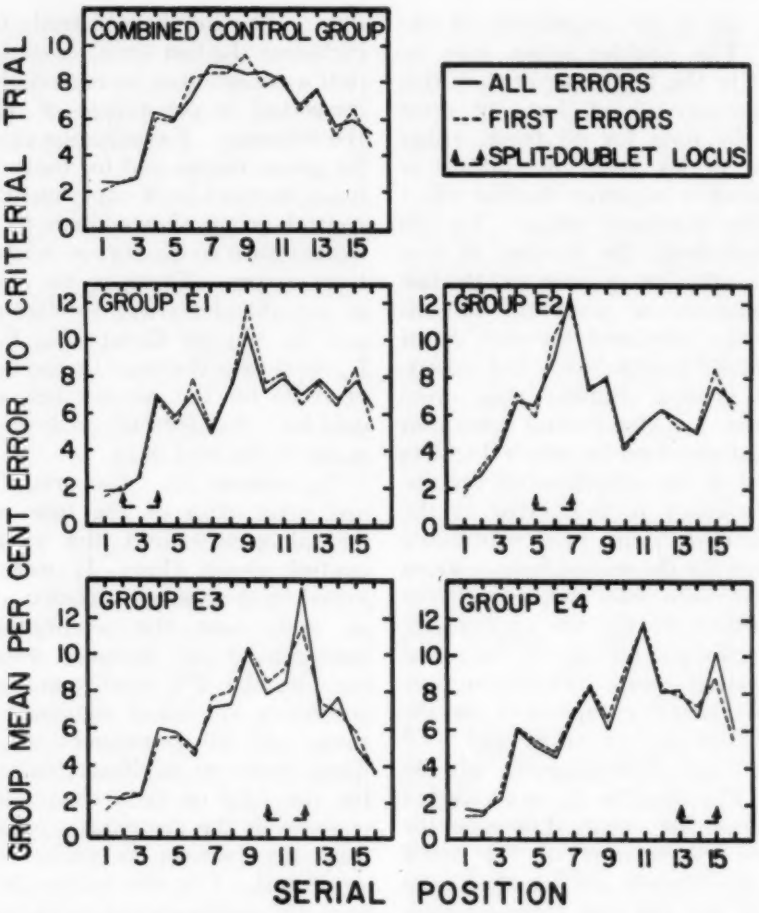


FIG. 3. Serial position curves in terms of percentage of total errors and percentage of first errors as a function of split-doublet locus.

error, and difference in percentage of first error between the experimental groups and the control group. The interaction of maze pattern and serial position was significant for each of these analyses. The differences produced in the form of the serial position error curve by the split-doublet are presented graphically in Fig. 3 and 4. The solid line curves represent total errors and the broken line curves represent first errors.

Examination of these curves shows no effect of the split-doublet for Positions 2–4 (E_1), an effect of increase in error at the final position of the split-doublet sequence for Groups E_2, E_3, and E_4, an increase in error for the second position of the split-doublet sequence for Groups E_2 and E_4, and no differences in curve form between the total error and first error measures. In addition, there are substantial differences between the experimental and control group serial position curves at positions other than the locus of the split-doublet.

The t test was used to establish fiducial limits at the 1% level for mean differences in percentage of total error between the experimental groups and the control group. The magnitude of positive difference in error (in favor of the experimental group) exceeds the fiducial limit for the third item of the split-doublet for Groups E_2, E_3, and E_4, and for the second item for Groups E_2 and E_4. The other positions at which the error differences exceed the fiducial limits are as follows: Group E_1, Position 7 (negative) and Positions 9, 14, and 15 (positive); Group E_2, Positions 10, 11, and 13 (negative), and Positions 3, 15, and 16 (positive); Group E_3, Positions 6, 13, and 16 (negative), and Positions 9 and 14 (positive); Groups E_4, Positions 6, 7, and 9 (negative), and Position 11 (positive). There are no consistent relations between these differences and specific positions or other features of maze pattern. They may be a complex function of the change in the maze pattern required to locate the split-doublet at the desired positions in the patterns of the experimental groups.

Further analysis of the split-doublet effect involved examination of the first-order contingent responses for the second and third split-doublet positions on the first trial. The results of χ^2 tests show a reduction in the frequency of response for a given number at the second position of the split-doublet sequence if it was the correct response for the first position. There are no differences in frequency of response at the third position of the split-doublet sequence as a function of the correct response for the first position. These results are comparable to those of Thompson, Voss, and Brogden (1957) and confirm their suggestion that the guessing habits of *Ss* during the discovery phase do not account for the split-doublet effect. Examination of the first error data by quarters of trials to the criterion shows that the split-doublet effect is present in the first error data in each one of the four quarters of trials and declines in

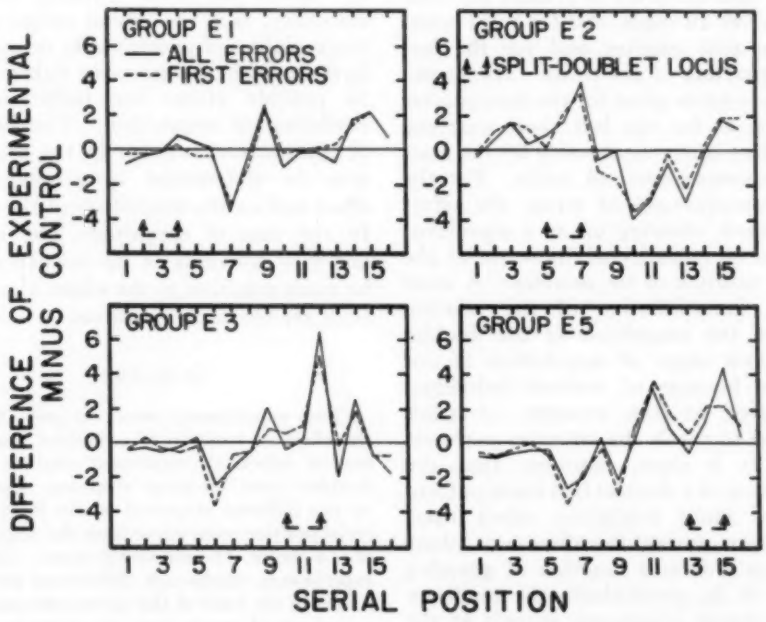


FIG. 4. Curves of differences between serial position curves for experimental minus control groups for percentage of total and first errors.

magnitude progressively from the first to the last quarter.

Discussion

The results of the two experiments establish a doublet effect and a split-doublet effect that are independent of each other. The doublet effect occurs at all locations in the maze pattern where tests for it were made, and there appears to be no difference in its magnitude as a function of the doublet locus. The split-doublet effect is differential in magnitude and in character as a function of its locus. The two phenomena are also different in that the doublet effect differs in the discovery and acquisition phases whereas the split-doublet effect does not. Both in the case of maze patterns involving doublets and split-doublets, there are other differences in form of the serial position error curves that may not be accounted for by the location of the doublet or split-doublet or by the interchanges necessary to their formation.

The doublet effect is evident for total errors over all trials, and for total errors for the first quarter and for the last three-quarters of the trials. Its magnitude is twice as great for the first quarter of trials as for the last three-quarters. The effect is not measurable by the first-error measure over all trials. For the last three-quarters of trials, the effect is reversed, showing up as a significant decrease in quantity of first-error at the second position of the doublet. A more precise determination of the relation between the magnitude of the doublet effect and stage of acquisition is not possible because of marked individual differences in the number of trials required to reach the criterion performance. It is clear, however, that the occurrence of a doublet in a maze pattern has an initial inhibitory effect upon acquisition during the discovery phase and that this is a function of guessing habits of *S*s, particularly the tendency not to repeat a response correct at the previous position. Once discovery is completed, the similarity of the items of the doublet facilitate further learning of the second item of the doublet. The latter hypothesis is supported only by results for the first-error measures, although the serial position curves constructed from the trial data support this in part by showing evidence of a decrease in trials for the second position of the doublet.

The split-doublet effect does not occur when the locus is early in the maze pattern and its nature for other locations is not completely consistent. It does involve an increase in error at the third position of the split-doublet sequence, and possibly also an increase for the second position of smaller magnitude than for the third position. There is no evidence of a difference in curve form for total errors and first errors, for successive stages of acquisition, nor of a relation between the guessing habits of *S*s and the split-doublet effect.

The results of these experiments suggest that in serial learning, variables of pattern may facilitate acquisition for the positions of their location or may inhibit it. When the serial learning involves discovery, as in the case of verbal mazes, the variables of pattern may or may not interact with the guessing habits of *S*s to provide either for facilitation or inhibition of acquisition. The location of the pattern variable in the sequence may be differential in producing an effect and on the magnitude of the effect. In the case of magnitude, the middle and late positions of the sequence may be more sensitive to the effect of pattern than are the early positions.

Summary

Two experiments were designed to test the effect of location of a doublet (two successive identical responses) and a split-doublet (two identical responses separated by one different response) on the form of the serial position error curve from the acquisition of a 4-choice, 16-unit verbal maze. In both experiments, significant differences were obtained in the form of the curves between each experimental group and the control group, and between experimental groups. In the case of the doublet, for each of the five loca-

tions, the effect was an increase in total error at the second position of the doublet. The effect did not occur for the first error data over all trials, but when the data for the first quarter and last three-quarters of trials were analyzed separately, an inhibitory effect was found for the first quarter and a facilitative effect was found for the last three-quarters of trials. The inhibitory effect is demonstrated to be a function of the guessing tendencies of the *Ss* not to repeat a response previously correct during the discovery phase. The split-doublet produced an increase in error for the third position for three of the four locations, and in two of these there was also an increase in error at the second position of the split-doublet. No relation was found between the guessing tendencies of *Ss* and the split-doublet phenomena, nor were there any differences in curve form between total errors and first errors or as a function of stage of acquisition.

References

Brogden, W. J., & Schmidt, R. E. Effect of number of choices per unit of a verbal maze on learning and serial position errors. *J. exp. Psychol.*, 1954, **47**, 235–240.

McCrary, J. W., Jr., & Hunter, W. S. Serial position curves in verbal learning. *Science*, 1953, **117**, 131–134.

Thompson, R. F., Voss, J. F., & Brogden, W. J. Effect of pattern variation upon verbal maze learning. *J. exp. Psychol.*, 1957, **54**, 253–258.

(Received March 4, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

RETROACTIVE INHIBITION OF CONNECTED DISCOURSE AS A FUNCTION OF PRACTICE LEVEL [1]

NORMAN J. SLAMECKA

University of Vermont

Retroactive inhibition (RI) varies inversely as a function of the degree of practice on original learning (OL) (McGeoch, 1929; Briggs, 1957), and, up to intermediate levels of interpolated learning (IL), directly as a function of the degree of practice on IL (McGeoch, 1932; Melton & Irwin, 1940; Underwood, 1945). All of these studies used nonsense syllables or other unconnected materials. However, there is doubt about the susceptibility of connected discourse to RI, particularly when rote retention measures are used (McGeoch & McKinney: 1934a, 1934b). We have been unable to find any evidence in the published literature that verbatim recall of prose is subject to significant RI. However, some preliminary work from this laboratory, using a group testing procedure, gave encouraging results (Slamecka, 1958), and the general importance of this problem warranted further investigation.

The present study was designed as a more precise test to determine whether varying degrees of practice would have the same effect upon the recall of connected discourse as they have upon unconnected materials. Specifically, it was hypothesized that the rote retention of a passage of prose is subject to RI, and that its recall is a direct function of amount of practice in OL and an inverse function of amount of practice in IL.

[1] This research was sponsored by National Science Foundation Grant G-6192. The author is indebted to Ernest N. Damianopoulos who assisted in the collection and analysis of the data.

METHOD

Subjects and materials.—Thirty-six students in general psychology at the University of Vermont served as *S*s as part of the requirements of the course. Each *S* served in four 40-min. sessions, the first being a practice period. Each session followed the RI paradigm, with OL, IL, and relearning of the OL passage by the method of serial anticipation. The materials were eight 20-word sentences, all taken from the same book source, modified slightly for length. Each word was exposed on the memory drum for 3 sec. Thus each trial took 60 sec., followed by a 6-sec. intertrial interval. The passages were grouped into four pairs, one of which was used for practice, and the others for the experiment proper. The first passage of each pair was for OL and the second for IL. The following three pairs of passages were used in the experiment proper:

(*a*) We must postulate that from strictly semantic points of vantage, most confusions in communication revolve about inadequate stipulation of meaning.

Communicators can exercise latitude in specifying meaning however they choose, provided that such definitions correspond somewhat closely to customary usage.

(*b*) So habitual become our expectations about symbols invariably possessing referents, that we tend unconsciously toward such assumptions concerning every word.

When words are utilized ostensibly as symbols although lacking appropriate referents, then perilous depths of semantic ignorance have been reached.

(*c*) However in most instances, linguists are predominantly concerned about determining the concrete denotations of discrete words functioning as proper names.

Much evidence regarding the inference of authentic meanings which are not stipulated, is derived from context analysis within the passage.

TABLE 1

MEANS AND *SD*s OF TOTAL CORRECT ANTICIPATIONS PER *S* DURING OL AND IL ACQUISITION

IL Trials	OL-2		OL-4		OL-8	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
	OL Acquisition					
0	18.00	5.00	40.33	12.88	105.17	21.16
4	17.42	6.38	40.17	8.69	107.17	22.12
8	16.33	5.57	44.50	11.30	97.33	22.24
	IL Acquisition					
4	32.58	8.6	37.08	15.66	43.83	11.24
8	89.00	21.27	109.25	20.22	114.75	26.10

Design and procedure.—A mixed factorial design (Lindquist, 1953, Type II) was used. One variable was the degree of OL practice (2, 4, and 8 trials), and the other was the degree of IL practice (0, 4, and 8 trials). Of the nine possible combinations of these levels of both variables, each *S* was given three, in such a manner that he had all levels of each variable but no repetitions of any level. The 36 *S*s provided 12 replications of the nine-cell matrix. Conditions were arranged so that the list pairs, sequence of sessions, and *S*s, were each counterbalanced with respect to the two main variables.

Each session began with two warm-up trials of color guessing. The colors appeared on the memory drum. Next, instructions for serial anticipation were given, followed by the appropriate number of OL and IL trials. In every case, 30 sec. elapsed between the end of OL and the start of IL, and 9.5 min. elapsed between the completion of OL and the OL relearning. To fill the temporal gap for the control (IL-0) and IL-4 conditions, a "rest activity" of color guessing was used in order to prevent rehearsal, while still preserving the set to respond to stimuli on the drum. Then, four relearning trials on the original passage were given, followed by three trials on the interpolated passage. Relearning of IL was done to motivate *S*s equally in the acquisition of OL and IL, thereby preventing laxness with respect to IL in later sessions.

TABLE 2

ANALYSES OF VARIANCE OF OL ACQUISITION AND RECALL

Source	*df*	Acquisition		Recall	
		MS	*F*	*MS*	*F*
Bet. *S*s	35	.031	.97	.069	1.07
OL × IL (b)	2	.016	.50	.142	2.20
Error (b)	33	.032		.065	
Within Ss	72	.165	17.85*	.114	2.93*
IL	2	.002	.20	1.088	28.06*
OL	2	5.629	608.67*	1.623	41.85*
OL × IL (w)	2	.006	1.25	.099	2.55
Error (w)	66	.009		.039	

* $P < .01$.

RESULTS AND DISCUSSION

Acquisition data for OL and IL appear in Table 1, in terms of the mean total correct anticipations per *S*. The OL data were subject to analysis of variance after a logarithmic transformation, and the results appear in the left half of Table 2. It is clear that increasing the OL trials, at every level of IL, produced a significant increase in the mean correct anticipations. Means within any IL level differed significantly from the adjacent means and there were no significant *F*s within any OL level. Analysis of the IL acquisition data indicated that increasing the

IL practice significantly increased the degree of acquisition. A tendency toward positive transfer was evident in that the IL means increased with increasing OL trials. Since this worked in a direction counter to the hypothesized effects of the main variables, the significance of the recall analysis, to follow, can only tend toward conservatism.

Table 3 shows the relearning of OL in terms of the mean total correct anticipations on each relearning trial. The recall data (first relearning trial) were analyzed after a log $(1 + X)$ transformation, and the relevant findings appear in the right half of Table 2. Both of the main variables were seen to be effective in determining recall. Within the levels used, recall as a function of OL practice approximates a linear function. An extrapolation through the OL-0 point gave a value of about two correct anticipations. It would be expected that even with no practice on such passages, *S* could guess a few words correctly by virtue of previous familiarity with the general language structure. Analyses within each OL level and within each IL level gave significant *F*s. Regardless of the degree of OL practice, the control group (IL-0) always recalled significantly more than the experimental groups (IL-4, IL-8). Thus, significant RI was obtained. It should be noted that the direction of effects of amounts of OL and IL was always as predicted even though all specific comparisons were not significant.

From the data of Table 3, relative RI $\left[\left(\frac{\text{Rest-Work}}{\text{Rest}}\right) \times 100\right]$, was calculated. The over-all RI was 43.0%. At IL-4 the relative RI for OL-2, 4, and 8, respectively, was 58.7%, 38.6%, and 24.3%. At IL-8 the values were 76.0%, 44.1%, and 34.7%, respectively. Thus, relative RI decreased as a function of increasing OL and decreasing IL practice.

The forgetting was rapidly dissipated within a few relearning trials. Table 3 shows that by the fourth relearning the differential effects of the interfering IL practice upon recall were largely dissipated, which

TABLE 3

MEANS AND *SD*s OF OL RELEARNING TRIALS

RL Trials	IL Trials	2 OL Trials		4 OL Trials		8 OL Trials	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
1	0	8.67	3.00	9.08	4.06	14.42	3.80
	4	3.58	3.16	5.58	2.53	10.92	3.10
	8	2.08	1.98	5.08	3.62	9.42	4.90
2	0	11.75	1.79	13.16	4.33	16.83	3.41
	4	10.16	3.53	12.50	2.66	16.25	2.83
	8	9.08	3.82	12.16	4.54	14.33	4.06
3	0	14.75	2.58	14.75	3.74	17.91	3.22
	4	11.25	3.48	14.50	2.30	18.50	1.65
	8	11.83	3.93	14.50	3.75	15.66	3.10
4	0	15.25	3.19	14.75	4.44	18.91	2.10
	4	13.16	5.48	15.66	1.79	19.33	1.18
	8	13.33	4.56	15.58	3.22	17.00	2.47

TABLE 4

CLASSIFICATION OF ABSOLUTE NUMBER OF ERRORS AT RECALL

Error Class	IL-0			IL-4			IL-8		
	OL 2	OL 4	OL 8	OL 2	OL 4	OL 8	OL 2	OL 4	OL 8
Omissions	100	68	37	141	129	73	174	145	92
Intralist	10	32	18	8	10	16	13	12	7
Interlist	0	0	0	7	12	10	5	5	12
Extraneous	24	25	11	32	21	6	23	15	15

is in accord with the usual findings based on unconnected materials.

The types of errors that occur during recall are often scrutinized for hypotheses concerning the processes underlying RI. A tabulation of all responses classed as errors is presented in Table 4. There are four categories of errors: (*a*) omissions, i.e., no response was made; (*b*) intralist intrusions, i.e., incorrectly placed words from the original passage; (*c*) interlist intrusions, i.e., words from the interpolated passage; and (*d*) extraneous intrusions, i.e., words found in neither passage. The use of connected discourse makes it impossible to categorize all intrusions unequivocally. For instance, a response which is classed as an intralist error is still a word in *S*'s language, and may have really originated as an extraneous intrusion. In addition, since connected discourse is serial in nature, errors cannot be subject to the S-R intrusion analysis traditionally done with the A-B, A-C paired-associate paradigm. In spite of these inherent limitations the data appear to provide tentative hints regarding the distribution of response tendencies.

The error data support two generalizations. First, that the better the recall, the less likely is an error to be an omission. This is the tendency even when omissions are stated as percentages of all errors made. Thus, under conditions of maximum recall (OL-8 and IL-0), only 56% of the errors are omissions, whereas under conditions of minimum recall (OL-2 and IL-8), 81% of all errors are so classed. Considering the two variables singly, there is still a trend toward decrease in the percentage of omissions as OL increases, and also as IL decreases. If omissions are considered as covert errors, and all other errors are considered as overt errors, then the data suggest a shift in the covert–overt error ratio paralleling changes in recall, reminiscent of the formulation of Thune and Underwood (1943).

The second generalization concerns the appearance of interlist intrusions. The frequency of these errors has been used as the index to the degree of competition attributed to the interfering material. The data suggest that such intrusions (whether listed in absolute or percentage terms) are more frequent when OL and IL acquisition is about equal. In terms of the percentage of all overt errors accounted for by this category, higher values are found at the OL-8-IL-8 (35%), and the OL-4-IL-4 (28%) levels, where the two lists are at about equal strength. Such a tendency has been reported for unconnected materials by Briggs (1957) and others. It appears to be consistent with a differentiation hypothesis (Gibson, 1941) whereby competing

responses of near equal strength are most likely to produce such intrusions, as compared to response systems more differentiated from each other by unequal amounts of practice.

With respect to the two hypotheses tested in this study, the following conclusions are drawn. First, that the rote retention of connected discourse has been shown to be subject to significant RI, and secondly, that such recall varies directly with degree of OL practice and inversely with degree of IL practice. These results serve to confirm the previous preliminary findings (Slamecka, 1958) and support the extension of generalizations hitherto based on the serial learning of unconnected materials to the serial learning of connected discourse.

Summary

In the RI paradigm, *Ss* learned passages of connected discourse by the anticipation method. Three levels of OL practice (2, 4, and 8 trials) and three levels of IL practice (0, 4, and 8 trials) were given, with appropriate counterbalancing in a mixed design. Results indicated that (*a*) OL and IL acquisition was a function of number of practice trials, (*b*) recall of prose was subject to significant RI, (*c*) recall varied directly with OL practice and inversely with IL practice, and (*d*) recall errors suggested a shift in the covert–overt error ratio paralleling recall, and an increase in interlist intrusions at points of near equal acquisition of the two passages. It is concluded that findings regarding degree of learning and RI, based on unconnected materials, can now be generalized to connected discourse.

References

Briggs, G. E. Retroactive inhibition as a function of the degree of original and interpolated learning. *J. exp. Psychol.*, 1957, **53**, 60–67.

Gibson, E. J. Retroactive inhibition as a function of the degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93–115.

Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

McGeoch, J. A. The influence of degree of learning upon retroactive inhibition. *Amer. J. Psychol.*, 1929, **41**, 252–262.

McGeoch, J. A. The influence of degree of interpolated learning upon retroactive inhibition. *Amer. J. Psychol.*, 1932, **44**, 695–708.

McGeoch, J. A., & McKinney, F. Retroactive inhibition in learning of poetry. *Amer. J. Psychol.*, 1934, **46**, 19–33. (a)

McGeoch, J. A., & McKinney, F. The susceptibility of prose to retroactive inhibition. *Amer. J. Psychol.*, 1934, **46**, 429–436. (b)

Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173–203.

Slamecka, N. J. Retention of connected discourse as a function of degree of learning. Paper presented at APA meetings, Washington, September, 1958.

Thune, L. E., & Underwood, B. J. Retroactive inhibition as a function of degree of interpolated learning. *J. exp. Psychol.*, 1943, **32**, 185–200.

Underwood, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, (**59** 3, Whole No. 273).

(Received March 7, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

MOTOR SKILL TRANSFER AS A FUNCTION OF INTERTASK INTERVAL AND PRETRANSFER TASK DIFFICULTY

GEDIMINAS NAMIKAS AND E. JAMES ARCHER

University of Wisconsin

In a recent pursuit-rotor study Lordahl and Archer (1958) examined the effects of transferring from one speed of rotation to a second speed of rotation. The *S*s practiced tracking a rotary-pursuit target at 40, 60, or 80 rpm on Day 1 and on Day 2 all *S*s practiced at 60 rpm. Substantial negative transfer obtained when there was a change in target speed.

The present study is concerned with the effect of variation of intertask interval upon this negative transfer. It was suggested in the Lordahl and Archer paper that, among other things, *S* learned a certain rhythm of responding. Conceivably the recovery from or forgetting of this rhythm could vary as a function of time since practice.

PROCEDURE

Subjects.—Women students from introductory psychology courses at the University of Wisconsin, who volunteered for the experiment, were used as *S*s. These *S*s received extra credit toward their course grade for serving in the experiment. Ten *S*s were assigned randomly to one of 12 conditions ($N = 120$). Four extra *S*s were run and their data discarded due to failure of either the tracking or recording apparatus. All *S*s were naive to the task and apparatus.

Apparatus and task.—The *S*s tracked a $\frac{3}{4}$-in. diameter, flush-mounted, brass target which was set 5 in. from the center of a pressed wood turntable. The turntable had a radius of 6 in. The tracking surface was 35 in. from the floor. The *S* tracked the target with a hinged stylus which prevented *S* from pressing on the handle and holding the stylus on target by force. The length of the stylus from the hinge to the point was 6 in. and the tip contained a steel ballbearing as illustrated in a previous study by Archer (1958).

The speed of the turntable was adjustable from about 10 to 85 rpm and a stroboscopic attachment allowed *E* to adjust the speed to within .5 rpm of the desired speeds. Performance of *S* was recorded on the Tracking Time Analyzer (TTA) which has been described in a previous paper (Archer, 1958).

Verbal signals of "Get set" and "Start" and "Stop" were given by *E* at the start and end of each trial respectively. The TTA and *E* were separated from *S* by a partial wall and *S*'s performance was not made known to her. The TTA was housed in a partially soundproof box, and a large noisy electric fan completed the masking of the clicking of its relays. The TTA was calibrated twice a day—morning and afternoon—to correct for fluctuations in line voltage.

Conditions.—All *S*s were given 20 training and 20 transfer trials. The trials were 30 sec. long with 60-sec. rest intervals interpolated between each. The *S*s trained for 20 trials on either the 40, 60, or 80 rpm speed and then had a rest interval of 2, 4, 8, or 16 min. All *S*s then transferred to the 60-rpm speed for 20 trials. Thus a 3×4 factorial design with speed of rotation orthogonal to intertask interval was used. During the intertask interval *E* changed the speed of rotation in the case of the 40- and 80-rpm groups and pretended to do so for the 60-rpm group by first changing the speed of rotation and then returning to the same speed. The three groups will hereafter be designated 40–60, 60–60, and 80–60.

Instructions.—All *S*s were given general instructions about the best way to track the target and a demonstration of such tracking by *E*. They were told to stay on target as long as possible. The *S*s were not told the number of trials they were going to have but the length of the trial and the rest interval were indicated. They were not informed about the longer rest interval, i.e., the intertask interval, beforehand, but after the training trials *E* told each *S* that the speed of rotation was going to be changed and that they were going to have a rest of a given duration. They were instructed to pick up the handle of the stylus on the signal "Get set" and then wait until the "Start" signal was given before starting to track the target.

During the intertrial and intertask intervals, after the signal "Stop," *Ss* were told to place the tip of the stylus in the center of the turntable and to sit down and look at magazines which were placed near by.

The *Ss* were not told anything about the recording apparatus. The answers to any questions, except those relating to the instructions, were deferred until the end of the experimental session. Instructions were tape recorded and were therefore identical for all *Ss*.

Results

The results are summarized in Fig. 1. This figure shows the mean percentage of time on target for all speed-interval combinations for both the training and transfer sessions.

Training session.—An extended Alexander trend analysis of variance (Grant, 1957) was performed on the training data. The intertask interval classification was included as a dummy variable to test the adequacy of the random assignment procedure and the assumption of initially equal subgroups within each speed condition. Since the test indicated that intertask interval was not an effective variable the assumption of equal subgroups seemed tenable.

The same analysis indicated that the speed of rotation was a highly effective variable. This agrees with the results of the Lordahl and Archer (1958) study. Time on target was found to be an inverse function of speed of rotation and the between-slopes and quadratics terms were highly significant (Fs = 221.63 and 63.44, respectively; $P < .01$ with 2 and 108 *df*). The significant linear component indicates that performance increases differentially as a function of

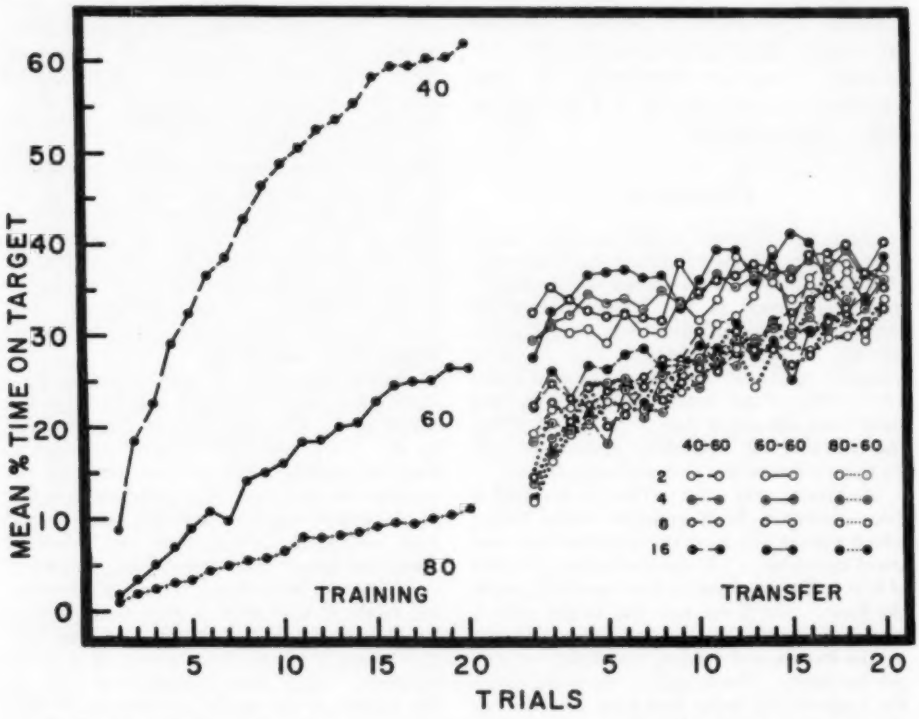


FIG. 1. Mean percentage of time on target for the training and transfer stages. Each data point in the training curves is based upon 40 cases; in the transfer stage each point is based upon 10 cases.

TABLE 1

A FREQUENCY DISTRIBUTION OF TIME CONTINUOUSLY ON TARGET IN THE FIRST FIVE TRIALS OF THE TRANSFER TASK

Training Task	Time Continuously On Target (In Sec.)										Total
	.00–.04	.05–.09	.10–.19	.20–.29	.30–.39	.40–.59	.60–.79	.80–.99	1.00–1.19	1.20–1.59	
40 rpm	2497	1525	2173	854	500	449	88	28	15	14	8143
60 rpm	2867	1849	2852	1137	792	692	230	73	50	33	10,575
80 rpm	3014	1600	1920	674	449	327	79	28	19	6	8116

Note.—The data are the total numbers of hits within each of the duration categories, with $N = 40$ for each training—task condition.

the speed of rotation, while the quadratic term shows a differential rate of increase in performance as a function of speed.

Transfer session.—An extended Alexander trend analysis of variance was performed on the transfer data also. Speed of rotation, during the training task, seems to have been the only variable to affect the transfer performance. There were significant differences between the means, linear, and quadratic components for speed ($F = 7.774$, $F = 13.473$, $P < .01$; $F = 3.550$, $P < .05$ with 2 and 108 *df*).

Because the total range of variation of the independent variable of intertask interval was relatively short (16 min.), it is conceivable that the influence of the intertask interval would be evident in only the first few trials on the transfer task. Accordingly, an Alexander trend analysis of variance was performed on the time-on-target data for the first five transfer trials only. Again, only the speed of the training task proved to have any significant effect. As in the analysis of the 20 transfer trials the between-group means, slopes, and quadratic components were found to be significant.

An extended trend analysis of variance was performed on the first five trials after transfer on the data from just the 40–60 and the 80–60 groups. No significant differences were found indicating that the above results were due to the differential performance of the 60–60 group as compared to the 40–60 and 80–60 groups.

Distribution of hits.—In Table 1 is presented the distribution of hits (frequency of times continuously on target) for the first five transfer trials. The numbers of hits for each duration were combined within each speed group without regard to the intertask interval. The superior performance of the 60–60 group is evident when the number of long-duration hits scored by this group is compared with the number scored by each of the other two groups (40–60 and 80–60). A chi square of 398.58 for 18 *df* ($P < .001$) was found for this table. These results are essentially the same as found by Lordahl and Archer (1958) and indicate that about the same transfer effects are found after an average of about 8 min. rest as compared to the 24-hr. interval studied by Lordahl and Archer. In order to complete both training and transfer sessions within a reasonable experimental period both of the sessions in the present study were made somewhat shorter than those in the Lordahl and Archer study. Therefore the magnitude of negative transfer after 8 min. and 24 hr. is not directly comparable.

Discussion

Speed of rotation.—The results substantiate the previous findings that variation in task difficulty, in terms of speed of rotation, has a very marked effect upon performance on both the training task and the transfer task.

In training, performance is found to vary inversely with the speed of rotation, i.e., better performance is found with slower speeds. The training condition was also found to have a very significant effect upon transfer performance. A major effect is due to a *change* in tracking speed. Since even in the first five trials after transfer, where the differences seemed to be greatest, there were no significant differences between the 40–60 and 80–60 groups, the above differences must have been due entirely to the better performance of the 60–60 group.

As suggested in the Lordahl and Archer study (1958) a possible hypothesis to explain these findings is that *S* learns a given set of motor components for performing a particular task. It is probable that a task of a given degree of difficulty emphasizes a particular set of components, which are acquired by *S*. With a change in difficulty, both training and transfer performance would be affected. For instance, an *S* in the 40-rpm group would probably learn a set of smooth, fairly slow, circular motions. These would tend to produce long duration hits and large time-on-target scores if he ever got on target. This *S* could be compared with one training in an 80-rpm condition. In this case a rapid jerky motion emphasizing quick adjustment to being off target would be acquired. Such a condition should produce a greater number of hits but most of them of short duration and probably with a smaller time on target score. On transferring to an intermediate difficulty condition, both of these *S*s would be under a handicap until they adjusted their performance to emphasize a particular combination of components most appropriate to this new condition. The rapidity of this adjustment would probably vary with the perceived similarity of the two tasks and the ability of *S* to realize the inappropriateness or inefficiency of his response.

Intertask interval.—Although *S* may learn the above hypothesized component skills appropriate to the training task difficulty, the performance of these component skills does not seem to vary differentially over a period of 16 min. In fact, the amount of negative transfer for any one group does not appear to vary as a function of these relatively short intertask intervals. Results indicate that the particular lengths of intertask intervals used in this study were completely ineffective in producing concomitant variations in performance.

Summary

The effects of three conditions of task difficulty were investigated on training and transfer performance on a rotary-pursuit task. The transfer was to the intermediate condition of difficulty used in training. The difficulty was varied by varying the speed of rotation in a rotary-pursuit task. Three speeds (40, 60, and 80 rpm) were used in training. All *S*s switched to 60 rpm in the transfer task.

Performance on transferring was also investigated as a function of the length of intertask interval. Four intertask intervals were used (2, 4, 8, and 16 min.).

The main findings were: (*a*) task difficulty (speed of rotation) was a very effective variable in both training and transfer performance, (*b*) performance was inversely related to task difficulty under training conditions, (*c*) transferring to the same speed gave the best performance, and (*d*) intertask interval did not affect transfer performance.

References

Archer, E. J. Effect of distribution of practice on a component skill of rotary pursuit tracking. *J. exp. Psychol.*, 1958, **56**, 427–436.

Grant, D. A. Analysis of variance in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.

Lordahl, D. S., & Archer, E. J. Transfer effects on a rotary pursuit task as a function of first-task difficulty. *J. exp. Psychol.*, 1958, **56**, 421–426.

(Received March 9, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

ANCHOR, CONTRAST, AND PARADOXICAL DISTANCE EFFECTS

HARRY HELSON
University of Texas

AND MYRTLE C. NASH
Southwestern at Memphis

The paradoxical distance effect (PDE) reported by Köhler and Wallach (1944) refers to the fact that test figures exposed in the vicinity of satiated or adapted visual areas suffer greater displacement the greater their distance, within limits, from the originally fixated figure. This finding is paradoxical because it appears to be in conflict not only with the almost universal physical law according to which effects diminish with distance from their source, but also with such basic psychological phenomena as lightness contrast and anchor effects where the rule is that stimuli near backgrounds or anchors suffer greater displacement than do stimuli farther away. While the theory of electronic currents advanced by Köhler and Wallach for figural aftereffects presumably accounts for PDE, it nevertheless leaves this interesting and important phenomenon in a class by itself. On the one hand it is reasonable to assume that lightness contrast and PDE are related since both result from interaction of retinal and/or central visual areas, and on the other it is not unreasonable to assume that phenomena in other sense modalities which are either materially or formally similar to PDE may be explicable in terms of a model which is independent of sense modality and which also accounts for background and anchor effects. While investigating effects of anchors on sets of weights at various distances from the anchoring stimuli, two types of shifts in judgments were found, one of which is analogous to classical contrast effects and the other to PDE. This study is concerned with these results and the possibility of dealing with them by means of a single theoretical model.

METHOD

Stimuli and procedure.—The stimuli consisted of 4-oz. medicine bottles filled with sand or mercury to make seven sets of weights as follows: Set I: 100, 150, 200, 250, and 300 gm.; Set II: 150, 200, 250, 300, and 350 gm.; Set III: 200, 250, 300, 350, and 400 gm.; Set IV: 250, 300, 350, 400, and 450 gm.; Set V: 300, 350, 400, 450, and 500 gm.; Set VI: 350, 400, 450, 500, 550 gm.; Set VII: 400, 450, 500, 550, and 600 gm. The sets of stimuli were presented either in ascending order, beginning with Set I and ending with Set VII, or in descending order, beginning with Set VII and ending with Set I. The weights in each set were judged in random order, and after five presentations the next set was judged without any break. To minimize interset effects, a control procedure was employed in which *Ss* judged Set VII immediately after Set I or Set I immediately after Set VII. Furthermore, the weights were first judged singly and one week or two weeks later were judged by all *Ss* following either a 90-gm. or a 900-gm. background stimulus. Counting the two orders of presentation (ascending and descending),

TABLE 1

JUDGMENTS OF SEVEN SETS OF WEIGHTS IN ASCENDING AND DESCENDING ORDERS BY METHOD OF SINGLE STIMULI (S.S.) AND FOLLOWING A 900-GM. BACKGROUND STIMULUS (B.S.) SHOWING DIFFERENTIAL EFFECTS OF THE BACKGROUND ON STIMULI WITHIN SETS (CONTRAST) AND BETWEEN SETS (PARADOXICAL DISTANCE EFFECT)

Ascending Series Order

Stimuli:	100	150	200	250	300	350	400	450	500	550	600	Total
S.S.:	20	41	61	75	86							
900 B.S.:	7	16	27	39	49							
Diff.:	−13	−25	−34	−36	−37							−145
S.S.:		34	48	62	72	80						
900 B.S.:		12	24	31	41	50						
Diff.:		−22	−24	−31	−31	−30						−138
S.S.:			39	58	70	83	89					
900 B.S.:			17	27	35	44	49					
Diff.:			−22	−31	−35	−39	−40					−167
S.S.:				48	64	72	81	90				
900 B.S.:				29	38	44	52	60				
Diff.:				−19	−26	−28	−29	−30				−132
S.S.:					50	62	76	82	89			
900 B.S.:					35	43	51	56	62			
Diff.:					−15	−19	−25	−26	−27			−112
S.S.:						55	65	77	86	88		
900 B.S.:						41	45	55	62	68		
Diff.:						−14	−20	−22	−24	−20		−100
S.S.:							58	68	80	83	88	
900 B.S.:							47	57	64	69	73	
Diff.:							−11	−11	−16	−14	−15	−67

Descending Series Order

S.S.:							44	56	69	77	80	
900 B.S.:							41	49	57	64	71	
Diff.:							−3	−7	−12	−13	−9	−44
S.S.:						40	54	66	74	83		
900 B.S.:						33	40	46	56	63		
Diff.:						−7	−14	−20	−18	−20		−79
S.S.:					33	47	60	71	80			
900 B.S.:					26	33	40	49	56			
Diff.:					−7	−14	−20	−22	−24			−87
S.S.:				30	45	56	70	80				
900 B.S.:				19	26	37	45	52				
Diff.:				−11	−19	−19	−25	−28				−102
S.S.:			27	43	56	69	80					
900 B.S.:			18	23	30	38	47					
Diff.:			−9	−20	−26	−31	−33					−119
S.S.:		23	37	51	68	77						
900 B.S.:		13	21	30	36	45						
Diff.:		−10	−16	−21	−32	−32						−111
S.S.:	22	36	51	66	75							
900 B.S.:	9	16	22	34	43							
Diff.:	−13	−20	−29	−32	−32							−126

TABLE 2

JUDGMENTS OF SEVEN SETS OF WEIGHTS IN ASCENDING AND DESCENDING ORDERS BY METHOD OF SINGLE STIMULI (S.S.) AND FOLLOWING A 90-GM. BACKGROUND STIMULUS (B.S.) SHOWING DIFFERENTIAL EFFECTS OF THE ANCHOR ON STIMULI WITHIN SETS (CONTRAST) AND BETWEEN SETS (PARADOXICAL DISTANCE EFFECT)

Ascending Series Order

Stimuli:	100	150	200	250	300	350	400	450	500	550	600	Total
S.S.:	20	37	54	70	75							
90 B.S.:	36	47	60	68	78							
Diff.:	+16	+10	+6	−2	+3							+33
S.S.:		35	50	63	72	82						
90 B.S.:		48	60	69	78	84						
Diff.:		+13	+10	+6	+6	+2						+37
S.S.:			46	56	68	78	83					
90 B.S.:			62	70	78	82	87					
Diff.:			+16	+14	+10	+4	+4					+48
S.S.:				48	60	70	76	84				
90 B.S.:				67	72	79	84	87				
Diff.:				+19	+12	+9	+8	+3				+51
S.S.:					57	65	71	78	84			
90 B.S.:					69	77	82	86	89			
Diff.:					+12	+12	+11	+8	+5			+48
S.S.:						60	68	76	80	86		
90 B.S.:						77	81	86	88	91		
Diff.:						+17	+13	+10	+8	+5		+53
S.S.:							65	73	79	82	88	
90 B.S.:							78	82	91	91	90	
Diff.:							+13	+9	+12	+9	+2	+45

Descending Series Order

S.S.:							39	50	61	74	77	
90 B.S.:							65	69	77	80	83	
Diff.:							+26	+19	+16	+6	+6	+73
S.S.:						37	47	57	70	75		
90 B.S.:						62	70	79	80	82		
Diff.:						+25	+23	+22	+10	+7		+87
S.S.:					34	46	55	65	70			
90 B.S.:					57	62	69	78	80			
Diff.:					+23	+16	+14	+13	+10			+76
S.S.:				32	43	51	64	74				
90 B.S.:				50	56	66	72	77				
Diff.:				+18	+13	+15	+8	+3				+57
S.S.:			26	39	54	62	71					
90 B.S.:			43	53	59	69	76					
Diff.:			+17	+14	+5	+7	+5					+48
S.S.:		24	35	50	60	68						
90 B.S.:		33	44	56	62	69						
Diff.:		+9	+9	+6	+2	+1						+27
S.S.:	16	29	45	54	64							
90 B.S.:	24	34	46	54	66							
Diff.:	+8	+5	+1	0	+2							+16

TABLE 3

SHIFTS IN ADAPTATION-LEVEL IN DIRECTION OF BACKGROUND STIMULI WITH GREATER MOVEMENT IN SETS FARTHER FROM VALUE OF BACKGROUND AS ANALOGUE OF PDE (ASCENDING ORDER)

Stimulus Sets	Adaptation-Level		Diff.	Adaptation-Level		Diff.
	Single Stimuli	900-Gm. Background		Single Stimuli	90-Gm. Background	
100–300	165	317	+152**	181	155	−26
150–350	202	357	+155**	198	159	−39
200–400	227	404	+177**	219	137	−82**
250–450	257	380	+123**	258	144	−114**
300–500	301	398	+97**	270	173	−97**
350–550	332	415	+83**	300	119	−181**
400–600	361	409	+48*	312	228	−84**

* $P < .05$.
** $P < .01$.

the two background stimuli, and the two modes of transition from lightest (heaviest) to heaviest (lightest) sets of stimuli (gradual and abrupt), it is seen that there were eight different conditions of judging. Different groups of 5 *Ss* each judged under each of the eight conditions, thus making a total of 40 *Ss* who served in this experiment.

Instructions.—The *Ss* were instructed as follows: "You will be presented with weights in succession. After lifting each weight you are to give a judgment by calling it very, very heavy; very heavy; heavy; medium heavy, medium, medium light; light; very light; or very, very light. If you need other categories such as extremely heavy or extremely light or terrifically light or terrifically heavy, you may use them. Be sure to judge each weight as it is and not as you think it should be." When the background weights were introduced, *Ss* were instructed to judge the second weight by itself, not as a comparison with the first weight. To acquaint *Ss* with stimuli and with the category scale, two practice rounds were given with the set of weights first presented.

Treatment of data.—The judgments given by *Ss* in terms of the rating scale were translated into numbers ranging from 10 for very, very light to 90 for very, very heavy in steps

TABLE 4

SHIFTS IN ADAPTATION-LEVEL IN DIRECTION OF BACKGROUND STIMULI WITH GREATER MOVEMENT IN SETS FARTHER FROM VALUE OF BACKGROUND AS ANALOGUE OF PDE (DESCENDING ORDER)

Stimulus Sets	Adaptation-Level		Diff.	Adaptation-Level		Diff.
	Single Stimuli	900-Gm. Background		Single Stimuli	90-Gm. Background	
400–600	419	452	+33	443	294	−149**
350–550	384	458	+74*	406	261	−145**
300–500	356	460	+104**	369	268	−101**
250–450	314	437	+123**	327	267	−60*
200–400	269	442	+173**	287	244	−43
150–350	232	404	+172**	251	226	−25
100–300	186	371	+185**	234	215	−19

* $P < .05$.
** $P < .01$.

of 10, with 50 for the medium category. In a very few cases some *Ss* employed categories more extreme than very, very light and very, very heavy, and these were assigned numerical ratings in accordance with their judged distance from the center of the rating scale. Values of adaptation-level as well as judgments of the stimuli expected according to theory were found from least squares fits to an equation formulated by Helson (1948) and also Helson and Himelstein (1955). The original ratings of the stimuli (in numerical terms), rather than theoretical values, are given in Tables 1 and 2, but the values of adaptation-level in Tables 3 and 4 are the fitted values. The influence of background (satiating) stimuli on judgments of the series stimuli appears in the differences between judgments of the stimuli presented singly and judgments following the background stimuli and also in the differences between the adaptation-levels for the sets of stimuli as judged in the two conditions. While the background effect is contaminated by carry-over effects from set to set, interset effects are also present when the stimuli are judged singly, and since the latter judgments serve as the norms by which shifts in judgment are gauged when anchors are introduced, interset influences may be regarded as common to both conditions in the comparison. The abrupt transitions from first to last sets and from last to first sets minimize interset effects and furnish a check on the results with continuous transitions from set to set.

Results

Two types of displacement.—The data presented in Fig. 1 and 2 and in Tables 1 and 2 reveal two distinct effects on judgments of series stimuli following introduction of the background stimuli. The first effect is the usual contrast effect in which stimuli nearer background suffer greater displacement than do stimuli farther from background. Thus the plots in Fig. 1 show that the heavier stimuli in both the heaviest and lightest sets of stimuli are displaced

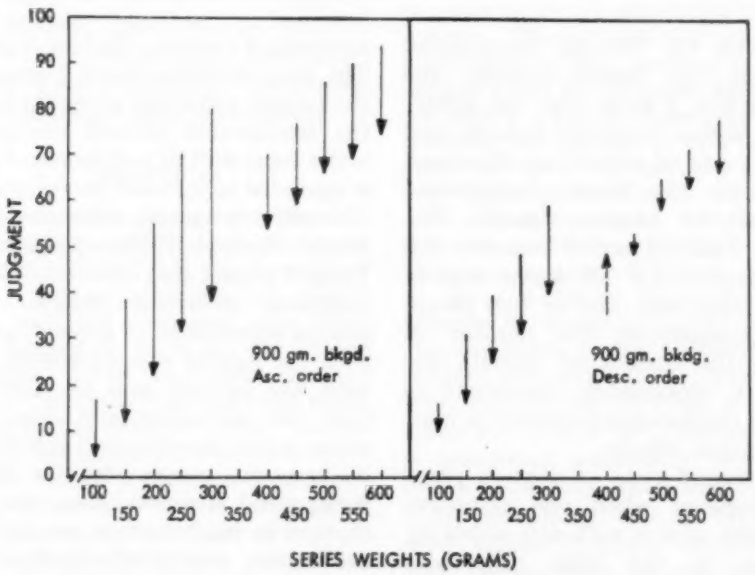


Fig. 1. Shifts in judgments of series stimuli following introduction of heavy background stimulus. The vectors represent direction and amount of shift and they reveal two types of displacements: (*a*) stimuli nearer background stimulus are displaced more than are stimuli farther from background (classical contrast and repulsion effect); (*b*) over-all displacements are greater in sets of stimuli farther from background stimulus (PDE).

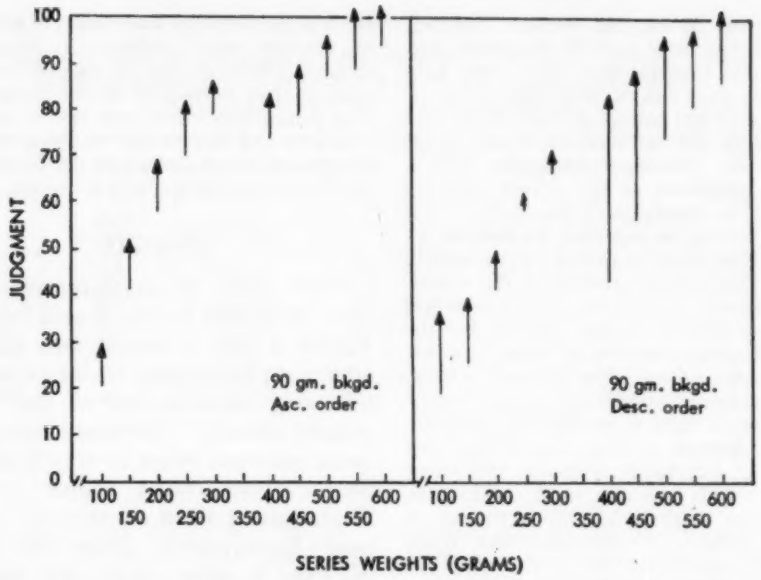


FIG. 2. Shifts in judgments of series stimuli following introduction of light background stimulus. Here the shifts are opposite in direction to those found with heavy background but the two types of displacements are nevertheless present.

more with the 900-gm. background than are the lighter stimuli; the plots in Fig. 2 show that the lighter stimuli within both the lightest and heaviest sets of stimuli are displaced more with the 90-gm. background than are the heavier stimuli. The data in Tables 1 and 2 bear out this generalization for all seven sets of stimuli with only one or two exceptions. Considering the number of *S*s and the variety of stimuli, the data are remarkably consistent in showing displacements similar to classical contrast effects.

The second type of shift in judgment appears when one compares the results of sets with sets either by reference to the total differences between judgments by method of single stimuli and with anchor stimuli (Tables 1 and 2), or by reference to the adaptation-levels in the two methods of judging (Tables 3 and 4). The data in Tables 1 and 2 show that the farther away any given set is from the background stimuli the greater is the total shift in judgments of series stimuli due to introduction of anchors. This effect is exactly analogous to the Köhler-Wallach PDE. Reference to Table 1 shows that the total shift in judgment with the 900-gm. background stimulus is, in general, greater with the lighter sets of stimuli than with the heavier sets of stimuli in both the ascending and descending orders while the data in Table 2 show the opposite occurs with the 90-gm. background stimulus: here the total changes in judgment are greater with the heavier sets of stimuli than with the lighter sets.

The changes in judgments of the series stimuli due to the two anchors or backgrounds reflect differences in

the adaptation-levels as seen from Tables 3 and 4. Thus the adaptation-level for the 100–300-gm. set changes from 165 gm. to 317 gm. when the 900-gm. anchor is introduced, a change of 152 gm., while the adaptation-level for the 400–600-gm. set changes from 361 to 409 gm. with the 900-gm. anchor, a change of only 48 gm. On the other hand the introduction of the 90-gm. anchor results in a change of only 26 gm. in the 100–300-gm. set but results in a change of 84 gm. in the 400–600-gm. set. The displacements of the neutral levels are not only greater, with sets farther from the background stimuli, they are, on the whole, also more significant statistically as shown by the *P* values associated with the differences between adaptation-levels for each of the sets of stimuli. The displacements analogous to PDE are also found in the abrupt transition from lightest (heaviest) to heaviest (lightest) sets of stimuli where interset effects are minimal so they cannot be attributed to prior stimulation.

Discussion

In the experiments reported here, two effects have been found: classical contrast effects and PDE, the former appearing when judgments of stimuli *within* sets are compared, the latter when judgments from set to set are compared. While the effects found here concern changes in felt weight as a result of introducing anchors much heavier or much lighter than the series weights, they are otherwise comparable to visual color and spatial contrasts and to PDE found in the spatial displacements of vision. Tactile-kinesthetic phenomena have been recognized by many writers, among them Gibson (1933), Köhler and Dinnerstein (1947), and Wertheimer and Leventhal (1958), as falling under the domain of satiation phenomena. Jaffe (1954) has demonstrated that visual extent may influence perceived tactile extent. The visual analogues to our findings are as follows: dark stimuli are lightened relatively more than are light stimuli on dark ground and light stimuli are darkened more than are dark stimuli on light ground (classical contrast), but all stimuli are lightened more on dark ground than they are on light ground while all stimuli are darkened more on light ground than they are on dark ground (PDE). Better still, if sets of grays are viewed on a dark gray ground and on a light gray ground simultaneously, it will be seen that the lighter grays on dark ground cannot be matched by any grays on light ground and the darker grays on light ground cannot be matched by any grays on the dark ground. Hence classical lightness contrast and a lightness PDE can both be observed if sets of gray stimuli are viewed on dark and light backgrounds simultaneously.

PDE depends on carry-over of a previous state of adaptation into the present. The old and new adaptation-levels form a new, intermediate level against which both old and new stimuli are judged. Köhler enunciated what is essentially our point of view when he said: "If there are currents which go with I-objects, there must also be currents corresponding to T-objects. . . . To a degree, we can tell how the obstructions established by the currents of given I-objects must distort the currents of T-objects" (1951, p. 212). In other words, as soon as the test figure appears, the satiated or adapted region must shift in the direction of the test figure and thus a new reference region is established somewhere between the region previously stimulated and the newly stimulated region. It immediately follows that test stimuli farther from the originally satiated region will be displaced more than nearer stimuli provided there is interaction between the old and new levels.

If this analysis, which covers both

classical contrast effects and PDE, is correct, then the adaptation-levels resulting from interaction of series and background stimuli should be displaced more the more the anchor stimuli differ from the mean of the series stimuli. This deduction follows from the definition of adaptation-level as a weighted geometric mean of series, background, and residual stimuli (Helson: 1947, 1948). It is immediately apparent that averaging 900 gm. with the 100–300-gm. series raises the level of this series more than averaging 900 gm. with the 400–600 series raises its level; conversely, the level of the 400–600-gm. series is lowered more than is the level of the 100–300-gm. series by averaging with the 90-gm. background stimulus. The values of adaptation-level calculated from the data of this study agree with predictions from the weighted mean definition of level regarding both direction and extent of change arising from introduction of background stimuli. A single theory therefore appears to be adequate to account for anchor or background and associated "contrast" effects as well as for the spatial displacements found in figural aftereffects.

There is, however, one respect in which we must limit the operation of the weighted mean definition of level. This definition applies only to the extent that there is interaction among stimuli and residual effects of stimulation and if interaction effects grow weaker, say with distance or some other factor, then account must be taken of this fact. Thus PDE is found only within certain limits, for if the test figure is too far from the region of previous stimulation there is no interaction and figural aftereffects may be diminished or even absent. A formula defining spatial levels would therefore require a distance function which would have to be determined for specific conditions of observation. A number of possibilities immediately present themselves and could be worked out within the general framework of the theory proposed here.

Summary

Eight groups of five *Ss* each judged seven sets of weights first by method of single stimuli and second after intervals of one or two weeks following either a very light or very heavy background or anchor stimulus. Two types of shifts in judgment due to introduction of the anchor stimuli were found: (*a*) within each of the sets, stimuli nearer the background stimuli were affected more than were stimuli farther from the background, an effect analogous to classical lightness contrast and the usual repulsion effects of anchor stimuli; (*b*) over-all shifts in judgment from set to set were greater in sets farther from than in sets nearer the background stimuli, an effect analogous to the paradoxical distance effect found in satiation experiments. It is shown that both classical contrast effects and paradoxical distance effects may be accounted for in terms of the adaptation-levels arising from interactions between series and background stimuli and between residuals from previous stimulation and present stimulation. Both the weighted mean definition of adaptation-level and the values of adaptation-level derived from the data of this experiment agree in predicting the direction and amounts of shift of judgment due to introduction of background stimuli. A single model would therefore seem to be applicable to classical contrast effects, repulsion effects of anchors, and paradoxical distance effects.

References

Gibson, J. J. Adaptation, after-effect and contrast in the perception of curved lines. *J. exp. Psychol.*, 1933, **16**, 1–31.

Helson, H. Adaptation-level as frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 1947, **60**, 1–29.

Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, **55**, 297–313.

Helson, H., & Himelstein, P. A short method for calculating the adaptation-level for absolute and comparative rating judgments. *Amer. J. Psychol.*, 1955, **68**, 631–637.

Jaffe, R. The influence of visual stimulation on kinesthetic figural after-effects. *Amer. J. Psychol.*, 1954, **67**, 668–676.

Köhler, W. Relational determination in perception. In L. A. Jeffress (Ed.),

Cerebral mechanisms in behavior. New York: Wiley, 1951. Pp. 200–243.

KÖHLER, W., & DINNERSTEIN, D. Figural after-effects in kinesthesis. In *Miscellanea Psychologica, Albert Michotte.* Louvain Inst. Sup. Philos., 1947. Pp. 196–220.

KÖHLER, W., & WALLACH, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Philos. Soc.*, 1944, **88**, 269–357.

WERTHEIMER, M., & LEVENTHAL, C. M. "Permanent" satiation phenomena with kinesthetic figural aftereffects. *J. exp. Psychol.*, 1958, **55**, 255–258.

(Received March 16, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

ELIMINATING A CONDITIONED GSR BY THE REDUCTION OF EXPERIMENTAL ANXIETY

ROBERT E. SILVERMAN

New York University

The GSR is especially sensitive to aversive motivational events. Experimental anxiety situations, in which aversive stimuli are presented, cause ordinarily neutral stimuli to elicit strong GSRs. This effect is generally termed sensitization or pseudoconditioning, because it does not depend upon the systematic pairing of a neutral stimulus with an aversive stimulus, as is the case in conditioning. Not only the direct presentation of aversive stimuli can produce sensitization, but also the induction of experimental anxiety by the threat of aversive events can do so. For example, Mowrer (1938) has shown that a neutral stimulus elicited GSRs when *S*s were fitted with electrodes, but that the same stimulus did not elicit GSRs when the electrodes were absent.

Since a sensitized response depends upon the presence or the threat of aversive stimuli, we would expect instructions which remove the threat to affect a sensitized GSR in much the same way as Mowrer's removal of the electrodes. However, a conditioned GSR might behave differently because conditioning is dependent, in part, upon the frequent pairings of the neutral and aversive stimulus. As a result of these pairings, the neutral or conditioned stimulus (CS) evokes a response which for a time is not dependent upon the presence of the aversive or unconditioned stimulus (UCS). The CS appears to act as a substitute for the UCS. If this is the case, then instructions about the absence of the UCS should be less effective with a conditioned response than they are with a sensitized response which depends upon the original UCS.

The purpose of the present study was to compare the effects of anxiety-lessening instructions on a conditioned and on a sensitized GSR. Anxiety-lessening instructions are defined as instructions which terminate the threat of aversive stimuli under the control of *E*. The results of two earlier studies indicated that a conditioned GSR (Cook & Harris, 1937) and a conditioned heart rate response (Notterman, Schoenfeld, & Bersh, 1952) can be influenced by instructions. However, in both of these studies the absence of a pseudoconditioning control group and the relatively long interstimulus intervals used make it uncertain whether the instructions were affecting a conditioned response or a sensitized response.

In order to determine whether anxiety-lessening instructions can affect a conditioned GSR as well as a sensitized GSR, three acquisition procedures were used to increase the probability of a GSR to a given CS. One group of *S*s was conditioned according to the classical Pavlovian paradigm with a CS-UCS interval of .5 sec. A second group was conditioned with a 6-sec. CS-UCS interval. This interval is considerably greater than the established intervals for classical autonomic conditioning, yet it is the interval used by some investigators to establish a heart rate CR (Notterman et al.,

1952; Zeaman & Wegner, 1954). A third group was a pseudoconditioning or sensitization group. For *S*s in this group, the CS and UCS were not paired; the UCS was presented randomly during the intertrial interval, never following the CS by less than 30 sec. nor preceding the CS by less than 10 sec.

Method

Subjects.—The *S*s were 60 male undergraduate students from the University College of Arts and Science and the College of Engineering of New York University.

Apparatus.—The CS was a 750-cps, 2-sec. tone delivered through earphones at approximately 20 db. The tones were on tape and the presentation of the tones served to program the intertrial intervals. A sound relay and two Hunter timers were used to control the interstimulus intervals and the duration of the CS and UCS. The UCS was a 6-sec. tetanizing shock to *S*'s left arm biceps.

The apparatus for measuring GSR was a modification of the unit described by Schlosberg (1954). It consisted of a 50-microampere meter in series with two quarter-sized silver-silver chloride-saline paste electrodes. The circuit was run by a single 1.5-v. flashlight cell which, with the aid of a potential divider and a calibrated resistor, permitted absolute conductance and GSR in terms of change in conductance to be read directly. In order to insure consistent electrode placement for each *S* and to maintain optimal contact during the experimental session, the electrodes were seated in rubber suction cups which in turn were seated in 4-in.-square wooden blocks. The whole unit was attached to *S*'s palms by means of ½-in. elastics.

The *S*s were seated in a totally darkened booth; an exhaust fan maintained air circulation and served as a masking noise. The *E* and the apparatus were in an adjoining booth.

Procedure.—The *S*s were assigned randomly to the experimental conditions. All *S*s, within each condition, were given an adaptation series consisting of 15 presentations of the CS. Just before beginning the adaptation series *S* was given the following instructions:

"This is an experiment to determine your reactions to tone and to electric shock. You will hear a series of tones through this headset which will soon be placed over your ears. A little later in the experiment, you will receive a series of electric shocks administered to your left arm; you will not be shocked through the electrodes I am now placing on your palms. The shocks you will receive will be strong enough for you to feel them, but will not be strong enough to do you any harm. Remember that they will just pass through your arm and not your body. Please try not to make any unnecessary movements. The apparatus is necessarily very sensitive, and your moving around too much would interfere with the recording."

The adaptation series was followed by the acquisition series consisting of 10 CS-UCS pairings. The interval between the CS and UCS depended upon the condition to which *S* was assigned. Three interstimulus intervals were used: a .5-sec. interval, a 6-sec. interval, and a random interval. In the latter condition the CS and UCS were never paired in that the CS was presented randomly at some point within the interval beginning 30 sec. after the UCS and ending 10 sec. before the next CS.

Following the acquisition series, half of the *S*s in each interstimulus condition were given the following instructions: "From now on you will not be shocked, so simply relax." These instructions were given by *E* who entered *S*'s booth. The remaining 10 *S*s in each of the conditions were given no instructions. All *S*s were given 15 extinction trials. The intertrial intervals for the adaptation, acquisition, and extinction series were programmed on tape and were the same for all *S*s. The intervals ranged from 50 to 75 sec.

Results and Discussion

The results were tabulated in terms of the mean GSR per *S* for each condition. These means were transformed into logarithms to reduce the marked heterogeneity of variance found in the raw data, and then means for each group were obtained. Table 1 shows the means and *SD*s of the transformed data for adaptation, acquisition, and extinction. An analysis of variance of the adaptation data indicated no significant differences among the groups, thus justifying the assumption that there were no systematic differences in response

TABLE 1

MEANS AND *SD*s OF THE TRANSFORMED GSRs FOR ADAPTATION, ACQUISITION, AND EXTINCTION

Extinction Conditions	Meas.	Adaptation			Acquisition			Extinction		
		Training Conditions			Training Conditions			Training Conditions		
		.5 Sec.	6 Sec.	C	.5 Sec.	6 Sec.	C	.5 Sec.	6 Sec.	C
Noninstructed	Mean	1.65	1.79	1.74	2.29	2.27	1.58	2.49	1.88	1.84
	SD	.18	.17	.20	.07	.10	.28	.09	.21	.19
Instructed	Mean	1.80	1.83	1.63	2.21	2.42	1.68	1.63	2.04	.94
	SD	.17	.14	.25	.13	.08	.28	.17	.14	.22

to the CS prior to the experimental treatments.

Table 2 presents the analysis of variance for the acquisition series. There were no significant differences during acquisition between the groups assigned to the different instruction conditions, since the instructions were not given until the extinction series. The differences among the three CS-UCS intervals during acquisition were significant. These differences were between the .5-sec. interval and the control, and between the 6-sec. interval and the control. The .5-sec. and the 6-sec. groups did not differ. Paired *t* tests indicated that the acquisition GSRs of both the .5- and 6-sec. groups were significantly different from the adaptation GSRs (.01 level).

An examination of Table 1 reveals that the increase in GSR amplitude in the 6-sec. group was short-lived in contrast to the increased GSRs found in the .5-sec. group. The mean of the uninstructed 6-sec. group falls back to the adaptation level during extinction, whereas the mean of the .5-sec. uninstructed group remains significantly higher than the adaptation mean (.01 level). This suggests that the .5-sec. group is conditioned, while the 6-sec. group is either weakly conditioned or merely sensitized to the impending shocks. The pseudoconditioning control group shows no reliable change in the mean GSR as a function of the acquisition procedures.

TABLE 2

ANALYSIS OF VARIANCE OF THE ACQUISITION DATA

Source	*df*	*MS*	*F*
Instructions	1	576.60	.16
CS-UCS Interval	2	30,379.65	8.43*
Interval × Instructions	2	673.55	.19
Within cells (error)	54	3,603.84	
Total	59		

* Significant at .01 level.

In order to control for the effects of the acquisition variables in analyzing the extinction data, the extinction means were adjusted in terms of the acquisition means and an analysis of covariance was performed. The analysis of covariance for the extinction series is presented in Table 3, where it can be seen that the instructions significantly affected the GSRs. The interaction calls attention to the fact that the instructions effect was not the same for each of the acquisition CS-UCS intervals. The .5-sec. group was markedly affected by the instructions, as was the control group. In both cases the GSRs were significantly reduced in amplitude (.01 level). However, the

TABLE 3

ANALYSIS OF COVARIANCE OF THE EXTINCTION DATA

Source	*df*	*MS*	*F*
Instructions	1	43,333.39	12.90*
CS-UCS Interval	2	23,667.11	7.05*
Interval × Instructions	2	17,477.93	5.20*
Within cells (error)	53	3,358.99	
Total	58		

* Significant at the .01 level.

6-sec. group did not appear to be so affected by the instructions. This latter finding is difficult to explain, in view of the fact that the control group responded to the instructions with diminished GSRs.

One possible interpretation is that for the .5-sec. and control groups, the instructions that no further shocks will be given tends to cancel the initial anxiety-arousing instructions given to all *Ss* at the outset of the experiment. If the initial instructions elicit an increase in GSR amplitude, as Mowrer's findings (1938) suggest, then overriding the initial instructions with later information should lower the GSR response. However, it is possible that the 6-sec. interval acquisition procedures produce different behavior, in that *Ss* develop a heightened state of GSR response which is no longer dependent upon the initial instructions. If this is the case, then the extinction instructions would not affect the GSR amplitude. In other words, the 6-sec. CS-UCS interval may act as a potent anxiety-arousing technique.

The fact that the 6-sec. interstimulus group did not evidence strong conditioning differs from the findings of Notterman et al. (1952), Zeaman et al. (1954), and Rodnick (1937), but these results are in agreement with those of White and Schlosberg (1952), as well as those of Moeller (1954). In both the Notterman and the Zeaman studies the CR was a measured change in heart rate during a 6-sec. interstimulus interval. Notterman et al. (1952) report a deceleration in rate as the CR, while the data of Zeaman et al. (1954) suggest that the form of the CR depends upon the point at which heart rate is measured. In both cases the possibility exists that the heart rate CR is an indirect effect of the preparatory skeletal responses occurring during the 6-sec. interval.

SUMMARY

Three groups of 20 *Ss* were conditioned to give a GSR to a 750-cps tone under different CS-UCS intervals. The intervals were: (*a*) .5 sec., (*b*) 6 sec., and (*c*) a random control interval in which CS and UCS were not paired. Half of the *Ss* in each group were given instructions prior to extinction that the UCS (shock) was not going to be given. The results indicated that the instructions affected both the .5-sec. group and the pseudoconditioning control group. The 6-sec. group was not affected by the instructions.

REFERENCES

COOK, S. W., & HARRIS, R. E. The verbal conditioning of the galvanic skin reflex. *J. exp. Psychol.*, 1937, **21**, 202–210.

MOELLER, G. The CS-UCS interval in GSR conditioning. *J. exp. Psychol.*, 1954, **48**, 162–166.

MOWRER, O. H. Preparatory set (expectancy)—a determinant in motivation and learning. *Psychol. Rev.*, 1938, **45**, 62–91.

NOTTERMAN, J. M., SCHOENFELD, W. N., & BERSH, P. J. A comparison of three extinction procedures following heart rate conditioning. *J. abnorm. soc. Psychol.*, 1952, **47**, 674–677.

RODNICK, E. H. Characteristics of delayed and trace conditioned responses. *J. exp. Psychol.*, 1937, **20**, 409–425.

SCHLOSBERG, H. Three dimensions of emotion. *Psychol. Rev.*, 1954, **61**, 81–88.

WHITE, C. T., & SCHLOSBERG, H. Degree of conditioning of the GSR as a function of the period of delay. *J. exp. Psychol.*, 1952, **43**, 357–362.

ZEAMAN, D., & WEGNER, N. The role of drive reduction in classical conditioning of an autonomically mediated response. *J. exp. Psychol.*, 1954, **48**, 349–354.

(Received March 16, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

PARTIAL AND CORRELATED REWARD IN ESCAPE LEARNING[1]

GORDON H. BOWER[2]

Yale University

Previous studies (Campbell & Kraeling, 1953; Bower, Fowler, & Trapold, 1959) have demonstrated that magnitude of reinforcement operates as a variable in instrumental escape learning. In those experiments rats were run in a straight alley in which the startbox and runway provided a continuous and intense electric shock, and the goalbox provided a continuous but weaker shock. It was found that running speed was faster the greater the reduction in shock intensity at the goal. Moreover, these performance effects were reversible; when the amount of shock reduction was changed, performance shifted to the level appropriate to the new amount of shock reduction.

Such studies have involved conditions in which *S* receives the same reward on every trial; however, of equal interest are those conditions in which the reward varies from trial to trial. Thus, a large reward (i.e., large amount of shock reduction) may be given on some trials and a small reward on others; or reward may be given immediately on some trials and delayed on others. Experiment I is an investigation of partial reinforcement which may be considered as a limiting case of variable magnitude of reward. Of further interest are those conditions in which variations in the reward are contingent upon variations in some characteristic of the response (e.g., its speed). In Exp. II, one condition of correlated reinforcement is investigated.

EXPERIMENT I

In this experiment the reward was varied independently of *S*'s performance. Two goal events were used, 100% or 0% reduction in shock intensity. Subjects received random sequences of these goal events, with the sequences differing in the proportion of 100% shock reduction trials. These conditions are somewhat analogous to partial reinforcement using, say, food and no food as the goal events.[3]

The primary interest in this experiment was to ascertain the relationship between speed and percentage of reinforcement when the goal events were given in a random sequence. Four groups of *S*s were trained with sequences corresponding to 100, 75, 50, or 25% reinforcement. A second question of interest was whether partial reinforcement in this escape situation causes slow adjustment of performance when the conditions are changed to 0% rewards, which is analogous to extinction.[3]

[1] This experiment was conducted during the writer's tenure on a predoctoral fellowship, MF-8585, granted by the National Institutes of Mental Health, Bethesda, Maryland.

[2] Now at Stanford University.

[3] Since a trial in the alley required eventual removal of *S* from the goalbox, the 0% shock reduction event may be described as a long delay of reward (removal from charged goalbox) instead of "nonreward," and as such might maintain performance above operant level. This description in terms of variable delay in no way changes the logic of the experiment.

Method

Subjects.—The *S*s were 20 male albino rats, about 100 days old. They were assigned randomly to four groups of five *S*s and were maintained with free access to water and Purina Lab Checkers.

Apparatus.—The apparatus is described in detail elsewhere (Bower et al., 1959). A 4-ft. straight alley with start and goal compartments was used. Two L-shaped strips of galvanized sheet metal, separated by a 1-in. floor gap, formed the two walls and halves of the floor. Two independent shock systems were connected to the sides of the alley; one shock system provided shock to the startbox and middle section of the alley, the other to the goalbox. The *S* received shock as long as its feet touched the two strips of sheet metal forming the halves of the floor. The 3 in. width of the alley prevented *S* from escaping shock by running along only one side of the alley.

Raising the start door operated a microswitch which closed the shock circuits to the alley and started a Standard Electric clock. The clock stopped when *S* interrupted a photobeam 2 in. inside the goalbox. Running times were recorded to the nearest .01 sec. and were converted to speed scores (1/time).

Design and procedure.—The shock intensity in the startbox and runway was always 250 v. The shock in the goalbox was either 250 or 0 v. Twenty acquisition trials were given. The four groups differed in the proportion (.25, .50, .75, or 1.00) of trials on which the goalbox shock was 0 v.; within these restrictions, random sequences of goal shocks were used. After the initial 20 trials, two groups (100% and 50%) were given 20 additional trials with the 250-v. goal shock present on every trial.

The *S* was placed in the start box and when it oriented toward the goal the start door was raised, turning on shock. When *S* entered the goalbox and interrupted the photobeam, the goalbox door was lowered and *S* was left in the goalbox for 20 sec. The trials were massed; immediately following removal from the goalbox, *S* was reinserted into the startbox for the next trial.

Results

The results in Fig. 1 show group mean speeds plotted in running averages of four training trials. Speed on the later acquisition trials is clearly related to percentage of reinforcement. Analyses of variance of mean speeds on Trials 13–20 demonstrate significant differences between the groups ($F = 12.16$, $df = 3,16$, $P < .001$). Component analysis of the group differences reveal that speed was a linear function of percentage of reinforcement (linear $F = 36.05$, $df = 1,16$, $P < .001$; residual $F < 1.00$). This linear relationship is shown graphically in Fig. 2, where a 0% control group from a previous experiment is also presented.

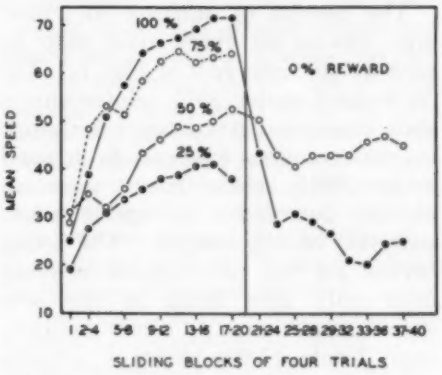


FIG. 1. Group mean speeds in sliding blocks of four trials. After Trial 20, the 100% and 50% groups were switched to a 0% reward condition.

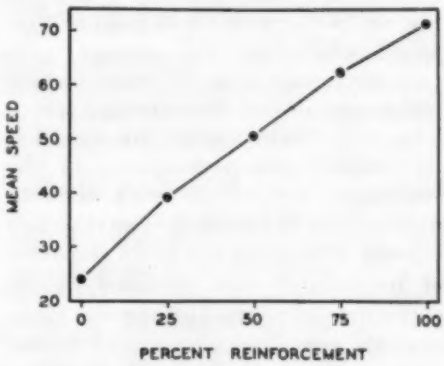


FIG. 2. Mean speed as a function of percentage or reinforcement. Each point is a group average speed over Trials 13–20.

The results of shifting to 250-v. goal shocks on every trial may be seen in the right half of Fig. 1. The *S*s trained under 50% reinforcement show considerable resistance to change in performance, whereas *S*s trained under 100% reinforcement show an abrupt decrement in speed when switched to 0% reward. The speed curves for the two conditions cross after only four trials at the new shock condition. A trend analysis over Trials 17–40 shows a reliable groups-by-trials interaction ($F = 7.05$, $df = 5,40$, $P < .001$), which demonstrates the difference in rate of shift.

Experiment II

In Exp. II, variations in reward were correlated with running speed, reward being given when *S* responded slowly. In the present study, the cutoff selected was a speed of 40, or running time of 2.5 sec. and 100% shock reduction was given for speeds below 40 and 0% shock reduction for speeds above 40. This condition is analogous to the DRL schedule studied extensively in the Skinner box (Ferster & Skinner, 1957).

Logan (1959) has proposed an equilibrium model for conditions of correlated reinforcement. The equilibrium performance is defined as that speed which, on the average, gets that percentage of reinforcement which will maintain just that average speed. The most direct test of this model is to compare the performance of the correlated reward *S*s with that of matched control mates. The matched control *S* receives the same sequence of rewards as that received by its experimental partner; however, these rewards are given irrespective of the control *S*'s speed. According to the equilibrium model, performance depends only upon the sequence of rewards and not upon the correlation of reward with speed. Because the sequence of rewards is equated, the equilibrium prediction is that correlated *S*s and their matched controls will perform similarly.

Method

Four pairs of rats were used, with the correlated and matched control *S*s of the pairs designated randomly. The apparatus and general procedure was the same as in Exp. I. The hair was shaved from *S*s' backs to provide better electrical contact. Each *S* received 50 massed trials. For an *S* in the correlated group, the 250-v. goal shock was turned off as soon as it broke the photobeam provided its running time exceeded 2.50 sec.; otherwise, the goal shock was left on. In either case, *S* was removed from the goalbox after 20 sec. The matched control procedure insured that a correlated *S* and its control partner were equated on the entire sequence of goal events.

Results

Group mean speeds in blocks of five trials are shown in Fig. 3. The curves start and rise together for 10 trials, but thereafter show progressive separation. The separation on Trials 41–50 was statistically significant ($F = 55.25$, $df = 1,6$, $P < .001$). The mean speed for correlated *S*s rises initially above the cutoff speed

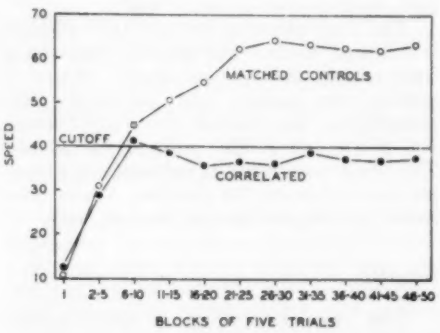


Fig. 3. Group mean speeds in blocks of five trials. The horizontal line represents the cutoff speed of 40.

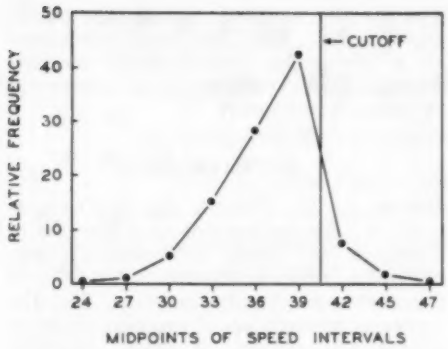


FIG. 4. Pooled speed distribution for the correlated group; plotted by intervals of three speed points.

of 40, but then declines to a relatively stable average speed just below the cutoff. The mean speeds over Trials 41–50 of the four *S*s in the correlated group were 37, 36, 36, and 38; their respective matched controls had mean speeds of 71, 65, 56 and 58. Apparently, the correlated condition reduced between-*S* variability to a large extent.

The pooled distribution of speeds over Trials 26–50 for correlated *S*s is shown in Fig. 4. A distinct mode appears in that speed interval just below the cutoff. Over the last 25 trials about 90% of their runs were rewarded.

DISCUSSION

The results of Exp. I show that speed is a linear function of the percentage of reinforcement; that is, the speed produced by random variations between the two rewards (100% and 0% shock reduction) is approximately the average of the speeds produced by the two rewards separately. Although this result was obtained with escape learning, it is consistent with the results of Logan[4] for variable magnitude and delay of food reinforcement. The result is contrary to the results obtained by Weinstock (1958) and Goodrich (1959) by varying percentage of food reward and nonreward. However, a number of procedural differences (e.g., number and massing of trials, nonrewards vs. long delays, the part of the response chain being measured, etc.) make serious comparisons difficult. That performance changes are slower following training with variable reward in this situation is consistent with the general findings for instrumental appetitive conditioning

The results of Exp. II clearly disconfirm the equilibrium prediction. Performance depends not only upon percentage of reinforcement but also upon the correlation of reward with speed. It should be noted that the performance of the uncorrelated, matched control *S*s is close to the speed expected on the basis of previous findings. Over-all, the percentage of reinforcement for the matched controls was about 80–85%. From Fig. 2, this percentage of reinforcement, when given uncorrelated with speed, would be expected to produce an average speed of about 64. This figure compares favorably with the average of 63 observed for the matched control group.

The fact that correlated *S*s learned to run slowly is consistent with the micromolar theory proposed by Logan (1956; 1959). In this approach, different speeds are treated as different responses, and the probability of a particular speed depends upon its reward relative to the rewards for competing speeds. The reinforcement dimension most relevant to the present results is the interval of reinforcement, defined as the time between the initiation of the response and the receipt of reward. The incentive for a particular response speed is assumed to be a decreasing function of its interval of reinforcement. For the present contingencies, this principle clearly favors speeds below 40 since they receive the reward immediately. Furthermore, although speeds of 38 and 20, for example, receive the same goal events, speeds of 38 have higher incentive because they have a shorter interval of reinforcement. In other

[4] Results reported in a forthcoming book, to be titled *Incentive*.

words, speeds of 20 are associated with 5 sec. of shock in the alley, whereas speeds of 38 are associated with only 2.65 sec. of alley shock. Thus, it is expected that *S* will learn to make predominately those response chains which minimize shock time, and the predicted mode of the speed distribution in Fig. 4 is at that speed interval just below the cutoff. Because the controlling contingencies "drive" each *S* into this optimal response region, the between-*S* differences are negligible.

Summary

Rats were trained to run down a straight alley to escape a 250-v. shock, the shock in the goalbox being either 0 or 250 v. In Exp. I the percentage of reinforcement was either 25, 50, 75, or 100, and it was found that speed was a linearly increasing function of this variable. When the 50% and 100% groups were switched to 0% reward, the 100% group extinguished faster than the 50% group.

In Exp. II reward was correlated with slow speeds, the goal shock being 0 v. for speeds below 40 and 250 v. otherwise. Matched control *S*s received the same sequence of goal events as that received by correlated *S*s, but reward was uncorrelated with the control *S*'s speed. It was found that *S*s in the correlated group learned to run at speeds just below the cutoff, receiving about 90% reward, whereas the matched controls ran significantly faster. This latter result was related to a micromolar theory which considers interval of reinforcement as one component of incentive motivation.

References

Bower, G. H., Fowler, H., & Trapold, M. A. Escape learning as a function of amount of shock reduction. *J. exp. Psychol.*, 1959, **58**, 482–484.

Campbell, B. A., & Kraeling, D. Response strength as a function of drive level and amount of drive reduction. *J. exp. Psychol.*, 1953, **45**, 97–101.

Ferster, C. B., & Skinner, B. F. *Schedules of reinforcement.* New York: Appleton-Century-Crofts, 1957.

Goodrich, K. P. Performance in different segments of an instrumental response chain as a function of reinforcement schedule. *J. exp. Psychol.*, 1959, **57**, 57–64.

Logan, F. A. A micromolar approach to behavior theory. *Psychol. Rev.*, 1956, **63**, 63–73.

Logan, F. A. The Hull-Spence approach. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. II. New York: McGraw-Hill, 1959.

Weinstock, S. Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *J. exp. Psychol.*, 1958, **56**, 151–158.

(Received March 20, 1959)

Journal of Experimental Psychology
Vol. 59, No. 2, 1960

THE EFFECT OF STIMULUS SIMILARITY ON AMOUNT OF CUE-POSITION PATTERNING IN DISCRIMINATION PROBLEMS

BARBARA NOTKIN WHITE[1] AND CHARLES C. SPIKER

Iowa Child Welfare Research Station

In the psychological literature, much attention has been paid to the problem of defining the effective stimuli in discrimination problems. Some of the discussion has been concerned with whether the response is conditioned to each of the individual components of a compound confronting *S*, or is conditioned to the compound per se. The early theory of discrimination behavior proposed by Spence (1936) and the later analysis of discrimination learning by Hull (1952) assumed that the response is conditioned to each component of a compound. Although Spence's (1952) theory has proved to be highly fruitful in explaining a number of phenomena discovered in the study of simultaneous discrimination learning, subsequent extensive study of the so-called successive discrimination problem led him to publish a theoretical analysis of this type of problem. He noted that his earlier theory was not applicable to the successive problem, since no component in the cue-position compounds of the successive problem is consistently reinforced more frequently than any other. He proposed the notion of cue-position patterning to account for the learning that occurs in the successive problem. In effect, he suggested that the response is conditioned to the compound consisting of a cue (e.g., black or white) and a position (e.g., left or right). Thus, *S* learns to discriminate between positive and negative compounds rather than between positive and negative components. Furthermore, Spence suggested that cue-position patterning occurs only when there is nondifferential reinforcement of the components (i.e., no one of the cue members of the compound is systematically reinforced more than the others).

The notion that response to patterns does occur under conditions where a component is differentially reinforced was directly investigated by Birch and Vandenberg (1952). They analyzed the cue-position compound (e.g., a black left alley) into three aspects—the spatial cue (S_L), the nonspatial cue (S_B), and a patterned stimulus (the compound) (S_{BL}). They noted that, in the simultaneous problem, both the nonspatial cues and the patterned cues are differentially reinforced; in the successive problem, on the other hand, only the patterned cues are differentially reinforced. Thus, only the patterned stimuli can serve as the basis of solution of the successive problem, whereas both types of cues might be used in the simultaneous problem. Accordingly, it is meaningful to ask, in connection with the simultaneous problem, to what degree the discrimination is determined by each of the two types of cues.

Birch and Vandenberg devised a convenient way of determining to what degree cue-position patterning develops in a simultaneous problem. Table 1 presents a schematic representation of a procedure suggested by them. In Problem 1, two simultaneous discrimination problems are used; however, only one spatial arrangement of each is presented, with the positive stimulus of the second setting in the position opposite from that of the positive stimulus in the first setting. In Problem 2, the test problem, the spatial arrangements of the two settings are reversed. In Problem 1, it will be noted that the left and right positions are equally

[1] Now at Michael Reese Hospital, Chicago.

TABLE 1

REPRESENTATION OF PROCEDURE FOR ISOLATING DEGREE OF CONDITIONING TO STIMULUS COMPONENTS AND TO PATTERN

Problem 1 (Training)	Setting 1	S_LS_A (+)	S_RS_B (−)
	Setting 2	S_LS_C (−)	S_RS_D (+)
Problem 2 (Testing)	Setting 1	S_LS_B ( )	S_RS_A ( )
	Setting 2	S_LS_D ( )	S_RS_C ( )

Note.—S_L and S_R refer to the left and right spatial positions; S_A and S_B refer to one pair of nonspatial cues (e.g., black and white); S_C and S_D refer to a second pair of nonspatial cues (e.g., red and green); (+) and (−) designate the reinforced and nonreinforced complexes of Problem 1; and ( ) refers to complexes which were neither reinforced nor nonreinforced (see text).

often reinforced and nonreinforced, whereas S_A and S_D are consistently reinforced and S_B and S_C are consistently nonreinforced.

If, in Problem 1, the response is differentially conditioned only to the nonspatial cue components, it should occur without disruption to the nonspatial cues of Problem 2. Thus, in the schema of Table 1, 100% response to S_A and S_D should be expected following mastery of Problem 1. It may be noted that the patterned stimuli of Problem 1 (S_{LA}, S_{RB}, S_{LC}, and S_{RD}) do not occur in Problem 2. Therefore, if the tendency for *S* to respond in Problem 1 has been determined entirely by differential conditioning to the patterned stimuli, he should respond equally often (i.e., by "chance") to the four stimulus compounds of Problem 2. Finally, if the response was differentially conditioned both to the nonspatial components and to the patterned stimuli of Problem 1, there should be a reduction in the tendency to choose S_A and S_D in Problem 2 due to the change in the patterned stimuli; but the choice of these stimuli should exceed 50% owing to the influence of the habits differentially established to the nonspatial cues. From these considerations, it is apparent that the amount of cue-position patterning that occurs in Problem 1 may be measured in Problem 2.

The results of the experiment by Birch and Vandenberg (1952) demonstrate rather convincingly that with rats cue-position patterning occurs in the simultaneous situation, despite the fact that the components of the cue-position compounds receive differential reinforcement.

One of the implications of the Birch and Vandenberg analysis is that the relative influence of the patterned stimuli in such a problem is determined by the relative ease with which differentiation can be achieved by means of the nonspatial components. Thus, in the simultaneous problem, involving highly distinctive nonspatial components within each setting, the role of the patterned stimuli should be negligible. At the other extreme is the successive problem in which the nonspatial components within each setting are identical. In the latter case, discrimination theoretically must be based entirely on differential habits established to the patterned stimuli. This suggests that the amount of cue-position patterning should be an increasing function of the degree of similarity between the nonspatial cues within each setting.

The purpose of the present experiment is to determine whether or not preschool children manifest cue-position patterning in a problem similar to that of Birch and Vandenberg, and to test the additional hypothesis that the amount of cue-position patterning is an increasing function of the similarity of the nonspatial cues within a given trial setting (i.e., the within-setting similarity). While the latter hypothesis is, in effect, a modification of Spence's theory, its confirmation within this procedure can at best be only suggestive, since the theory is explicitly restricted to inarticulate organisms.

METHOD

Experimental design.—In order to test these hypotheses, three groups of *Ss* were run, each with a different degree of similarity between the two nonspatial cues presented on each trial (Table 2). For Group I, the stimuli were dissimilar; for Group II, similar; and for Group III, identical. The test situa-

tions were the same as the corresponding training situations for each group, except that the spatial position of each cue stimulus was reversed. For each group, differential responding to the components was reflected in *S*'s tendency to select, on each test trial, the previously reinforced nonspatial cue. Conversely, response to patterns was indicated by *S*'s tendency to select, within each setting, the stimulus in the position previously occupied by the positive stimulus of the training problem.

On the assumption that all discrimination problems involve some patterning, and that the relative effectiveness of the patterned stimuli is an increasing function of within-setting stimulus similarity, it was expected that the choice of the stimulus in the previously reinforced position would be very low for the group trained on dissimilar stimuli (Group I), very high for the group in which the stimuli were identical (Group III), and somewhere in between for the group with quite similar stimuli (Group II).

Apparatus.—The stimuli were eight 3 × 3-in. pieces of colored paper backed by heavy white plastic. On every trial, the two stimuli to be used were fitted into the front faces of two 3½-in. cubical wooden boxes mounted 6 in. apart on a 17-in. tray. These were presented to the child on top of a stand 12 in. high. The boxes were separately hinged to the tray on the side away from the stimulus slot so that they could be lifted by *S* to obtain (on reinforced trials) a marble located on a groove in the tray. Each marble so won was placed in a small plastic container. The entire apparatus, with the exception of the marble container, was painted flat black.

Subjects.—The *S*s were 42 children from the Iowa Child Welfare Research Station preschool laboratories, their ages ranging from 4 to 5½ yr. The data of an additional 14 *S*s were discarded; one because of incorrect instructions, the other 13 because of failure to learn the problem in a reasonable number of trials. Of these, five were from Group I, four from Group II, and five from Group III. The *S*s were assigned by a random procedure to each of the three treatment groups; each treatment group finally containing 14 *S*s.

Procedure.—The design of the experiment is presented in Table 2. It can be seen that the only essential difference among the training problems of the three groups lay in the degree of similarity of the stimuli presented on a trial. In the case of the stimuli for Group II, the discriminability between the red and orange-red and between the white and light gray were subjectively judged by several independent observers to be about equivalent. It should also be noticed that the correct stimuli and the sides on which they were placed were the same for all groups.

All *S*s within any group were trained to a criterion of four consecutive correct responses on the type of trial involving red as the reinforced stimulus, and then received training to the same criterion on the second type of trial. Once both criteria were met, further training continued as follows: the different types of trials were alternated according to a predetermined random order which was repeated three times in 36 trials. When necessary, one or two blocks of six trials were interspersed to break up position or alternating habits that appeared to be forming. When *S* reached a subcriterion of 5/6 correct responses in a block of six trials, he was presented with the criterion block of six trials, the order of which differed from the order used in training. If he did not meet

TABLE 2

SCHEMATIC REPRESENTATION OF THE EXPERIMENT

Group	Setting	Train (Problem 1)	Test (Problem 2)
I	1	$S_L S_{red}$ (+) $S_R S_{yellow}$ (−)	$S_L S_{yellow}$ ( ) $S_R S_{red}$ ( )
	2	$S_L S_{black}$ (−) $S_R S_{white}$ (+)	$S_L S_{white}$ ( ) $S_R S_{black}$ ( )
II	1	$S_L S_{red}$ (+) $S_R S_{orange}$ (−)	$S_L S_{orange}$ ( ) $S_R S_{red}$ ( )
	2	$S_L S_{gray}$ (−) $S_R S_{white}$ (+)	$S_L S_{white}$ ( ) $S_R S_{gray}$ ( )
III	1	$S_L S_{red}$ (+) $S_R S_{red}$ (−)	$S_L S_{red}$ ( ) $S_R S_{red}$ ( )
	2	$S_L S_{white}$ (−) $S_R S_{white}$ (+)	$S_L S_{white}$ ( ) $S_R S_{white}$ ( )

Note.—S_L and S_R refer to the left and right spatial positions; (+) and (−) designate the reinforced and nonreinforced complexes of the training problem; and ( ) refers to complexes which were neither reinforced nor nonreinforced. The stimuli appeared in the spatial positions indicated above.

this criterion of all six correct, training was continued, and a second opportunity on the criterion block allowed. The final performance criterion was the same for all groups.

For each of the three groups, the test situation consisted of 10 presentations of the same stimuli used during training, with one difference—the side on which a given stimulus was placed was changed (see Table 2). The order of presentation of the different types of trials was the same as that used for the latter part of training. However, one-half of the *S*s in each group started the series with one type of trial (Subgroup A), the second half with the other type (Subgroup B); then each followed through the same predetermined random sequence of changes from one type of trial to the next.

Before entering the experimental room, each *S* was allowed to choose one of a selection of dime store toys to serve as a prize. He was told that he would be permitted to play a game in which he could win enough marbles to buy the toy. On each trial, if *S* selected correctly, he found a marble under the box. The noncorrection procedure was used throughout.

During the test trials, no reinforcement was given, so that the tendencies to respond remained relatively unchanged after the training series. In an attempt to avoid the extinctive effects of nonreinforcement, the following method was used: *E* informed *S* that she would remove the marble for him, that she would keep track of the number of marbles he won, and that when the game was over, she would show him how many marbles he had won. During this phase, *S* was asked only to point to the box which he wished to select, and *E* did not inform him whether or not his choice was correct.

Results

Examination of the training data (Table 3) reveals little difference among the three groups in the mean number of trials to criterion. A simple analysis of variance comparing the groups gave a nonsignificant F ($F < 1.0$). Greater difficulty of the successive problem would not necessarily be expected in this paradigm as the successive problem involved only two nonspatial cues while the other two problems each involved four.

TABLE 3

Trials to Criterion in Training Problem and Responses to Previously Reinforced Position

Group	Trials to Criterion		Patterned Responses	
	Mean	*SD*	Mean	*SD*
I	26.71	14.52	2.36	2.50
II	26.29	11.95	6.50	3.64
III	24.57	17.70	9.36	1.59

Considering the test trials, the responses to the previously reinforced position were tabulated for each treatment group, keeping separate the data of the A and B subgroups (which counterbalanced for the type of test trial first received). Since t tests showed no differences between the A and B subgroups within any treatment, these were collapsed in the final analysis of the data. Table 3 gives the means and *SD*s for each treatment group over all 10 test trials. As was expected, the mean number of responses to the previously reinforced side increased as the similarity of the within-pair stimuli increased (i.e., from Group I to Group III). A test suggested by Cochran and Cox (Snedecor, 1948) was used to analyze the results, since the within-groups variance, owing to the almost complete uniformity in the performance of *S*s in Group III, differed among treatment groups. An analysis over all 10 test trials indicated that Groups I and II differed at the .01 level, and that the difference between Groups II and III was also significant ($.05 > P > .02$).[2]

In Groups I, II, and III, the numbers of *S*s responding to position on the first test trial were 3, 9, and 14

[2] A Mann-Whitney U test was also performed on the data, yielding essentially the same results.

respectively. A chi square test comparing these scores gave a value of 18.38, significant at the .001 level for 2 *df*.

The analyses of all test trials and of the first test trial, considered together, seem to indicate that the effects of training appeared early in the test series and remained relatively constant throughout all 10 test trials.

Discussion

The results of the present study support the contention that the effectiveness of the patterned stimuli in any trial setting (i.e., the degree of patterning) increases as the similarity of the stimuli for that setting increases. In Groups I and II, if the pattern were not an effective cue, *S*s would have continued to respond consistently to the previously reinforced nonspatial stimulus when these stimuli were shifted in the test problem. In the case of Group III, unless response to patterns had occurred, Problem 1 would have been insoluble; since the test problem was the same as the training problem, patterning was again revealed by selection of the originally reinforced pattern. Hence, in each group, response to patterns was manifested in each of the two stimulus settings of the test problem by selection of the stimulus presented in the same position as that of the previously reinforced stimulus. This manifestation was greatest in Group III, and least in Group I.

The data for Group III may be used to determine whether or not the introduction of the novel suspension of reinforcement in the test situation resulted in a general disruption of performance. Since Problem 1 was the same as Problem 2 for this group, any marked change in performance between the two problems could be interpreted as the result of changes in the conditions under which *S*s performed. The performance of Group III, however, showed very little change from Problem 1 to Problem 2. This fact, together with the fact that the analysis of the first test trial leads to the same conclusion as does the analysis of all 10 test trials, indicates that the introduction of the test did not disrupt performance.

The possibility remains, and must be dealt with, that some of the results reported in support of the hypothesis are artifacts of the experimental design and the analyses used.

Firstly, it is possible that the increasing amounts of cue-position patterning with increasing within-setting stimulus similarity reflected not individual behavior, but rather group tendencies, and were the results of averaging over *S*s, some of whom showed response to nonspatial cues only, some of whom responded to patterns only. If this were so, then increase in patterning with increase in similarity would be due to an increase in the proportion of *S*s who showed all (as opposed to no) patterning. That such a factor does not entirely account for the results is suggested by the following tabulations: in Group I, only two *S*s gave as many as 6 responses to patterns, while nine gave 0 or 1. In Group III, only two *S*s showed less than 10 such responses, and neither of these two showed less than 5; in the case of Group II, three *S*s showed between 0 and 2 responses to patterns, five gave between 3 and 7, and six *S*s gave between 8 and 10.

Another possibility directly concerns Group II where an intermediate number of responses to patterns was obtained. To rule out the possibility of this having been obtained by the majority of *S*s responding consistently to one position on all 10 trials, two analyses were made. The number of *S*s giving between 8 and 10 responses to one side was three, while the majority (11) showed between 5 and 7 responses to any one side (hence, between 5 and 3 to the other). This corresponds with Groups I and III where the numbers of *S*s (out of 14 in each case) responding 8 or more times to one position were 4 and 1 respectively. Also, if *S*s responded consistently to one position in all 10 test trials, they would of necessity show

"response to patterns" consistently to one of the two settings and no patterning-type responses to the other setting. Hence, another way of tackling the same problem is to correlate the numbers of patterned responses in Setting 1 with those in Setting 2. For the 14 *Ss* in Group II, the Pearson r was .672 ($P < .01$), indicating that those *Ss* who patterned low or high in one setting tended to respond similarly in the second setting. (The same analysis over all 42 *Ss* yielded a correlation of .635 ($P < .001$), but this is probably inflated by the preponderance of 5's for both settings in Group III).

Thus it would seem that the results obtained are indicative of increasing responses to cue-position patterns with increasing stimulus similarity, and are not simply an artifact of averaging or of *Ss* adopting a position preference on the test trials.

Some of the work done by Bitterman and his associates is relevant to the thesis presented here. Teas and Bitterman (1952) reported data which correspond exactly to the predictions generated by the hypothesis of this paper. In addition, a statement by Bitterman and Wodinsky (1953, p. 375), although phrased in different terminology and couched in a different theoretical framework, appears to be, at least in part, a statement of the hypothesis proposed here.

Summary

This experiment studied the effects of increasing the similarity of the within-pair stimuli in a discrimination problem on the amount of patterned responding in a nonreinforced test series. Three groups of preschool children were trained to respond to a red stimulus on the left side on one type of trial and to a white stimulus on the right side on a second type of trial in a discrimination situation involving the choice of one or the other of two boxes fronted by colored cue stimuli. For one group, the nonreinforced cue stimuli were quite dissimilar to the rewarded stimuli with which they were paired. For another, the dissimilarity between each pair was less, and for the last, the two stimuli presented on a trial were identical. It was found that in the subsequent test series, which involved the training stimuli in reversed spatial position, the number of responses to the side previously correct for a given stimulus setting (i.e., patterned responding) increased significantly as the similarity increased.

References

Birch, D., & Vandenberg, V. The necessary conditions for cue-position patterning. *J. exper. Psychol.*, 1955, **50**, 391–396.

Bitterman, M. E., & Wodinsky, J. Simultaneous and successive discrimination. *Psychol. Rev.*, 1953, **60**, 371–376.

Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

Snedecor, G. W. *Statistical methods.* Ames: Iowa State Coll. Press, 1948.

Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.

Spence, K. W. The nature of response in discrimination learning. *Psychol. Rev.*, 1952, **59**, 89–93.

Teas, D. C., & Bitterman, M. E. Perceptual organization in the rat. *Psychol. Rev.*, 1952, **59**, 130–140.

(Received April 7, 1959)

ZEITSCHRIFT FÜR
EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Journal of the German Psychological Association

Editors for Germany

PROF. DR. PH. LERSCH, München; PROF. DR. F. SANDER, Bonn

Editor for Austria
PROF. DR. H. ROHRACHER, Vienna

Editor for Switzerland
PROF. DR. R. MEILI, Bern

Redaction
DR. C. J. HOGREFE, Göttingen

Volume 6, No. 2

HANS JUERGEN EYSENCK. **The Maudsley Personality Inventory as a Determinant of Neurotic Tendency and Extraversion**

WALLACE A. RUSSEL and OSKAR R. MESECK. **The Influence of Association on Word Memory in German, French, and English**

FRANZ WEINERT. **The Influence of Practice on the Task-Oriented Set of Ten-Year-Old Children**

HEINZ SCHMIDTKE. **Analysis of Job Completion with Special Reference to Rest Periods**

EVA GRUHNWALD and EBERHARD ULICH. **On the Use of Work Experiments with the Psychologically Handicapped with Special Reference to an Investigation of Deaf Youths**

WERNER TRAXEL and GUNHILD WREDE. **Changes in Psychogalvanic Skin Response as Affected by Musical Selections**

ELFI ZEBLEWSKI. **The Applicability of Mittenecker's Perseveration Test to Personality Diagnosis**

RAINER ROEHLER. **On the Possibility of Quantitative Comparative Measurements with Mierke's Determination Apparatus**

KARL MIERKE. **Tests of Driving Ability by Use of the Kiel Determination Apparatus**

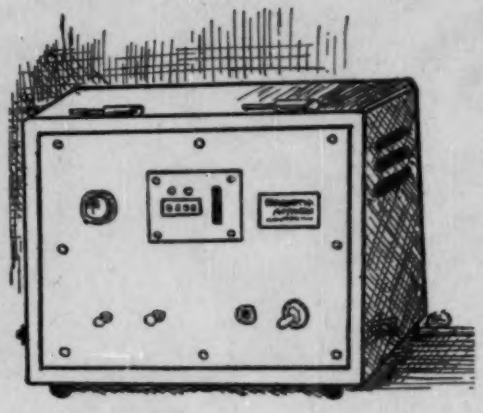
All articles are provided with summaries in English and French.

The journal appears quarterly and contains 160 pages per issue.
Subscription: DM 64.—per volume.

VERLAG FÜR PSYCHOLOGIE—DR. C. J. HOGREFE—GÖTTINGEN
Brentanoweg 10, Germany

MARIETTA APPARATUS COMPANY

Presents the new 24-15 Low Current Impulse Counter.



FEATURING

- **Counting rates up to 30 per second.**
- **Accumulation to 9,999 with manual reset.**
- **A handsome, durable case with detachable dust cover.**
- **Operation on 110 Volt, 60 Cycle alternating current.**
- **Contact current less than 0.001 ampere at 6 Volts.**
- **Visual indication of counts by flashing light.**
- **No batteries to replace**
- **Price $85.00**

This unit is ideally suited for the psychological laboratory as it will count any events that can be made to open and close an electrical circuit. The low current feature eliminates arcing of contacts. This increases the reliability and adaptability of the counter and also decreases distractions during psychological tests.

Catalogue #60, listing additional Psychological Laboratory Equipment is available upon request.

MARIETTA APPARATUS CO.

Marietta, Ohio
U. S. A.